Molecular modeling of solute partitioning in micellar systems

Vom Promotionsausschuss der
Technischen Universität Hamburg
zur Erlangung des akademischen Grades
Doktor-Ingenieurin (Dr.-Ing.)

genehmigte Dissertation

von
Denitsa Yordanova

aus
Shumen, Bulgarien

2020

1. Gutachterin: Prof. Dr.-Ing. Irina Smirnova

2. Gutachter: Prof. Dr. Dr.h.c. Frerich Keil

Vorsitzender des Prüfungsausschusses: Prof. Dr.-Ing. Michael Schlüter

Tag der mündlichen Prüfung: 17. Januar 2020

Danksagung

Diese Arbeit entstand im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Institut für Thermische Verfahrenstechnik der TU Hamburg. Ich möchte insbesondere Frau Professor Dr.-Ing. Irina Smirnova für die Unterstützung während dieser Zeit danken. Mein Dank gilt für das entgegengebrachte Vertrauen in meine Arbeit, die stets offene Tür, die gute Atmosphäre und die angenehme Zusammenarbeit.

Herrn Prof. Frerich Keil danke ich für die Begutachtung der Dissertation sowie Herrn Prof. Michael Schlüter für die freundliche Übernahme des Prüfungsvorsitzes. Herr Prof. Frerich Keil und Dr. Sven Jakobtorweihen haben mich von der ersten molekülardynamischen Simulation während der Bachelorarbeit bis zur Fertigstellung meiner Dissertation bei meiner wissenschaftlichen Arbeit begleitet, dafür möchte ich mich herzlich bedanken. Dr. Sven Jakobtorweihen danke ich für die fachliche Betreuung dieser Arbeit, für das kritische Durchlesen der Dissertation und die hilfreichen Kommentare sowie für die angenehme und produktive Zusammenarbeit.

Danken möchte ich auch allen Studenten, welche durch ihre Abschlussarbeiten zum Gelingen dieser Arbeit beigetragen haben. Weiterhin bedanke ich mich bei allen Kollegen am Institut für die freundschaftliche und hilfsbereite Arbeitsatmosphäre. Insbesondere möchte ich meinen Kollegen aus der mizellaren Gruppe, Eric, Ralena, Thomas und Simon für die enge und kontruktive Zusammenarbeit, für die wertvollen Diskussionen und für die exzellente gemeinsame Zeit am Institut danken.

Vielen Dank auch an Steffi für Ihre unerschöpfliche Hilfsbereitschaft.

Mein Dank gilt darüber hinaus meinen Freunden und meinem Partner Vasil für die Unterstützung auch in stressigen Zeiten. Ganz besonders danke ich meiner Familie für das unendliche Vertrauen und die Unterstützung, auf die ich mich immer verlassen kann.

Contents

Abstract **v**

1 Introduction **1**

2 Fundamentals **5**

2.1 Micelles and liposomes 5

2.2 Thermodynamics of micellar systems 9

2.2.1 Thermodynamics of micellization 9

2.2.2 Formation of micelles in mixed surfactant solutions . . . 12

2.2.3 Partition equilibria in micellar systems 13

2.3 Molecular dynamics simulations 18

2.3.1 Force fields 20

2.3.2 Position restraints 26

2.3.3 Free energy calculations 27

2.4 COSMO-RS and COSMOmic 30

2.4.1 The Conductor-like Screening Model for Realistic Solvation 31

2.4.2 COSMOmic for anisotropic systems 33

3 State of the art **37**

3.1 Molecular dynamics simulations of micellization 37

3.2 Free energy profiles in micelles 39

3.3 Partition equilibria prediction of neutral solutes in micelles . . . 39

3.4 Micelles composition of mixed surfactant systems 41

3.5 Partition equilibria of ionizable solutes in membranes and micelles 42

3.6 Aim of this work 45

4 Materials and methods **47**

4.1 Experimental determination of partition coefficients with micellar liquid chromatography 47

4.2 Analysis of micelles from MD simulations 47

4.2.1 Definition of micelles 47

4.2.2 Shape analysis 48

4.3 Averaged atomic distributions for COSMOmic 49

4.4 COSMOmic for free energy profiles in micelles 51
4.5 Thermodynamic cycle and position dependent pK_a 52
4.6 Error calculation 54

5 Modeling of Triton X-114 and Triton X-100 micelles 55
5.1 Optimization of force field parameters 55
5.1.1 Computational details 55
5.1.2 Parameter optimization overview 56
5.2 Self-assembly in aqueous solution and partition equilibria 64
5.2.1 Simulation details 64
5.2.2 Self-assembly: Influence of concentration and temperature 65
5.2.3 Prediction of partition equilibria: Influence of micelle size and shape 69
5.3 Summary 72

6 Free energy profiles and partition equilibria in various micelles: Comparison between MD simulations, COSMOmic and experiments 73
6.1 Micellar liquid chromatography for CTAB micelles 73
6.2 Simulation details 76
6.2.1 Preassembled micelles 76
6.2.2 Umbrella sampling protocol 77
6.2.3 COSMO-RS and COSMOmic 78
6.3 Calculated free energy profiles and partition coefficients 79
6.3.1 Nonionic surfactants 81
6.3.2 Zwitterionic surfactants 82
6.3.3 Anionic surfactants 84
6.3.4 Cationic surfactants 86
6.4 Summary 96

7 Micellization and partition equilibria in mixed nonionic/ionic micellar systems 99
7.1 TX114/SDS and TX114/CTAB mixed micelles 99
7.1.1 Simulations details 99
7.1.2 Formation of mixed micelles: Self-assembly MD simulations 100
7.1.3 Partition equilibria: Comparison between experiments and predictions 102

7.2 Brij35/CTAB mixed micelles 103
7.2.1 Simulation details 104
7.2.2 Partition coefficients of neutral solutes 104
7.2.3 Partition coefficients of ionized solute 108
7.3 Summary 111

8 Micelle/water compared to liposome/water partition coefficients 113
8.1 Simulation details 114
8.2 Partition coefficients in DMPC/water and HePC/water: COSMOmic versus experiments 114
8.3 Summary 118

9 Conclusions 119

Bibliography 123

Abstract

This work focuses on the modeling of solute partitioning in various micelles using predictive atomistic models. The combination of all-atom molecular dynamics (MD) simulations and the COSMOmic model provides a detailed and at the same time computationally effective method to describe partition equilibria in micellar systems. Therefore, new insights on the partition behavior of solutes in different types of single and mixed surfactant systems are gained.

Force field parameters, which are an essential part of every MD simulation are optimized and validated for nonionic surfactants of the Triton X series to enable their modeling with this technique. The self-assembly process of micelle formation in nonionic and mixed nonionic/ionic surfactant systems is studied with MD simulations. Predicted partition coefficients of neutral solutes in nonionic, zwitterionic, anionic, cationic and mixed nonionic/ionic micelles are in good agreement with experimental data.

Since no predictive models have been established to predict the partition behavior of charged compounds in micelles, in this work the prediction quality of the methods is evaluated with a special focus on ionizable molecules. In order to enable this evaluation, micelle/water partition coefficients of ionized solutes are determined experimentally. To gain detailed information on the localization of ionizable solutes within different types of micelles, free energy profiles, which cannot be obtained experimentally, are calculated with COSMOmic and umbrella sampling MD. Furthermore, an approach to improve the prediction quality of COSMOmic for negatively charged molecules in mixed nonionic/cationic micelles is proposed. Based on free energy profiles and a thermodynamic cycle, a position dependent pK_a within the micelles is estimated. Therefore, the protolytic equilibrium not only in water but also in the micelle is taken into account.

This work provides insights on the partition behavior of neutral and charged solutes in micellar systems on an atomistic level. By considering various solutes and surfactant systems, the applicability of MD simulations with the COSMOmic model for screening studies is evaluated.

1 Introduction

Surfactants are chemical compounds of particular significance due to their broad spectrum of applications in life sciences and in the chemical industry. A surfactant is an amphiphilic molecule consisting of a hydrophilic head group and a hydrophobic tail group. Surfactants are classified according to the nature and charge of their hydrophilic heads: nonionic, anionic, cationic and zwitterionic surfactants. In aqueous solutions at concentrations higher than their critical micelle concentration (CMC) surfactants self-assemble into aggregates (micelles), whereby their hydrophilic part is directed towards the aqueous phase, while their hydrophobic tails point to the center of the micelle.

The water solubility of hydrophobic compounds can often be enhanced with micelles due to the favorable partitioning of solutes in the micelles compared to water.[1–3] Because of this property known as solubilization, micellar systems find various applications. In cosmetics, water-insoluble compounds are dissolved using the solubilization in micelles.[4–7] Micellar solubilization is used for extraction and purification of membrane proteins.[8–13] In pharmacy, surfactants are often added to drugs to improve the dispersion within the human body and are used as drug delivery agents.[4,14,15] In nonionic surfactant solutions, a phase separation is observed, when the temperature is increased above the cloud point temperature (CPT). In this case the surfactant solution separates into two phases: a surfactant-rich phase and an aqueous-rich phase, whereby hydrophobic compounds are favorably solubilized within the micelles, hence they are concentrated in the surfactant-rich phase. Due to this feature of nonionic surfactants, they find application in environmental analytical chemistry to enrich and extract nanoparticles[16] and in chemical engineering applications for separation processes known as cloud point extractions (CPE).[17–24] For applications including proteins or living organisms it is important that the CPT is not much higher than a specific temperature. In this context, the nonionic surfactant Triton X-114 is widely used for membrane protein purifications and cloud point extractions.[17–23,25] Zwitterionic micelles are used to mimic biological membranes due to their structural similarity to phospholipids and their easier preparation compared to liposomes.[26–28] In many applications, mixtures of surfactants are used, because they have improved characteristics compared to single surfactants solutions.[29–31] For example, mixed micelles from block copolymers are promising vehicles for drug delivery due to improved physical stability and enhanced loading capacities compared to conventional polymeric micelles for drug delivery.[32] In separation processes, the extraction efficiency can be enhanced due to synergetic effects when mixtures of nonionic and ionic

surfactants are used.[33,34]

In pharmaceutical applications, the partition behavior of charged solutes is of particular interest, since many drug molecules are ionized to some extend at physiological pHs.[35,36] The knowledge of the drug affinity to the lipophilic environment at different pH values is crucial to mimic physiological processes.[37] In this context, mixed micelles are promising drug delivery agents, as they can enhance the solubilization of drugs and therefore to contribute to increased bioavailability, controlled drug release and targeted delivery.[38–40] The partition equilibria of ionizable solutes is of importance in extraction processes as well. By adjusting the pH value, the surfactant types and the micelle compositions, the partition behavior of solutes can be tailored for a specific application.[41]

In order to evaluate and compare the partition behavior of a solute in different micellar systems, a quantitative property is required. The partition equilibria of the solubilized compound between the micelle and the aqueous bulk phase can be quantified with a partition coefficient. Partition coefficients can be obtained experimentally using for example micellar liquid chromatography (MLC),[27,42] micellar enhanced ultrafiltration (MEUF),[42,43] solubility determination[44] and derivative spectroscopy.[26,28] However, depending on the application many surfactant-solute combinations are possible and the optimal combination should be identified. Therefore, theoretical methods to predict the partition behavior in micellar systems are of prime importance, especially screening tools are desirable. There are few approaches in the literature, which allow the prediction of micelle/water partition coefficients. One group are models based on linear correlations between physicochemical properties and partition coefficients (e.g., property-property relations (PPR)[45], linear solvation energy relations (LSERs)[46], quantitative structure-activity relationships (QSARs)[47,48]). However, these models require physicochemical properties of the system of interest. For screening studies of new compounds predictive models are desirable. In this context, a promising tool is the "conductor-like screening model for realistic solvation" (COSMO-RS).[49,50] In this thermodynamic model, the micelles are treated as a homogeneous surfactant phase in equilibrium with the aqueous bulk phase (pseudo phase approach).[51] However, in reality there are anisotropic systems. For more precise and more detailed (e.g., solute partitioning at different regions within the micelle) investigations of the partition behavior, models which take into account the anisotropy of the system should be used.

Molecular modeling encompasses all computational techniques used to predict or mimic the behavior of molecules. These methods are an essential part of the modern life sciences and are used in the fields of computational chemistry, computer aided drug design and bioinformatics. All-atom molecular dynam-

ics (MD) simulations are a computational technique to predict the behavior of molecules at an atomistic level. Therefore, MD simulations are suitable for molecular modeling of anisotropic systems. In MD simulations, the interactions between atoms and molecules are defined in force fields. Hence, to simulate a system of interest force field parameters for the present molecules have to be available. Moreover, the prediction quality of a MD simulation is strongly dependent on the force field parameters. Based on the behavior of the atoms, processes such as self-assembly and partition equilibria can be modeled with MD simulations. In presence of solute molecules, their favorable partitioning within the simulated system can be estimated based on free energy profiles. Free energy profiles, which are not accessible by experimental methods, can provide insights into the localization of solutes within micelles. However, the calculation of free energy profiles directly from MD simulations is computationally demanding and therefore only reasonable for systems of special interest. Another approach is the combination of MD simulations with the COSMOmic model,[52] which is an extension of COSMO-RS for anisotropic systems such as micelles and membranes. In this case, MD simulations are performed to obtain micelle structures, which are necessary for the calculations with COSMOmic. However, once a micelle structure is obtained, the COSMOmic calculations are computationally efficient, therefore this approach is suitable for screening studies.

The aim of this work is to evaluate the combination of MD simulations with the COSMOmic model for partition equilibria prediction of neutral and charged solutes in various single and mixed micellar systems. Thereby, different types of surfactants are considered including nonionic, zwitterionic, anionic, cationic as well as mixtures of nonionic/anionic and nonionic/cationic surfactants. The partition coefficients prediction is evaluated based on comparison with experimental data. In addition, partition behavior is also calculated directly with MD simulations using the umbrella sampling technique. Hence, in this work the three-dimensional micelle structure is taken into account, in contrast to the literature, where most theoretical studies of solute partitioning in micellar systems apply the pseudo phase approach, which neglects the anisotropy of the micelles.

This thesis is organized as follows. In the next chapter, the fundamentals of micellar systems and the applied methods are summarized. In chapter 3, previous studies, relevant for this work are outlined and the aims of this work are derived. Methods, which are applied or proposed in this work are explained in chapter 4. The results of the work are presented in chapters 5 to 8. Chapter 5 contains force field parameters optimization of Triton X molecules, modeling of

self-assembly and partition equilibria in Triton X-114 and Triton X-100 micellar systems. In chapter 6, free energy profiles of neutral and charged solutes in four types of micelles are represented. Self-assembly MD simulations, as well as partition equilibria predictions in mixed nonionic/ionic micellar systems are introduced in chapter 7. In chapter 8, calculated with COSMOmic micelle/water and liposome/water partition coefficients are compared to experimental data. Finally, the findings of this work are concluded in chapter 9.

2 Fundamentals

2.1 Micelles and liposomes

Molecules containing both polar and nonpolar groups are called amphiphiles. They can self-organize into ordered structures due to hydrophilic and lipophilic interactions. In these self-assembled aggregates, the monomers are orientated such that the polar part of the molecules is in contact with a polar environment and encloses the nonpolar part, or reversed.[53] Well known amphiphiles are surfactants and polar lipids. Surfactants can be classified according to the charge of their hydrophilic group as: nonionic, zwitterionic, anionic and cationic. Sodium dodecyl sulphate (SDS) and cetyltrimethylammonium bromide (CTAB) are the most commonly used anionic and cationic surfactants, respectively. Conventional nonionic surfactants, such as the Triton X series and Brij series (C_iE_j) consist of a poly(ethylene oxide) chain with different chain length as a polar group.[54]

Depending on the molecular structure of the amphiphile, its concentration in the solution and physical properties of the system, such as temperature and ionic strenght, they are able to form various structures and phases. Amphiphiles with one hydrophobic tail tend to pack into structures with a positive intrinsic curvature (micelles), where cylindrical molecules with two hydrophobic tails organize into flat bilayers.[55] At high concentrations in aqueous solutions, amphiphiles form macroscopic liquid-crystalline phases (e.g., lamellar, hexagonal and cubic phases), which upon dilution can be dispersed into relatively stable colloidal particles (e.g., micelles and liposomes).[56] Some examples are shown in Figure 2.1.

Spherical micelles are the simplest structures and form spontaneously in aqueous solution, when the surfactant concentration exceeds a threshold value, known as critical micelle concentration (CMC). The CMC is specific for each surfactant. Micelles have an anisotropic water distribution within their structure, whereas the water content decreases from the surface to the micelle core, which is mostly water-free.[57] Schematic illustrations of cationic and nonionic spherical micelle structures are shown in Figure 2.2. In ionic micelles, the hydrophobic core is surrounded by the Stern layer, which is a shell of hydrophilic head groups and contains counterions, considered to be bound to the micelle. The Gouy-Chapman electrical layer is a diffuse layer, which is surrounding the Stern layer and contains counterions as well. Its thickness depends on the ionic strenght of the solution and decreases in presence of electrolytes.[57,58]

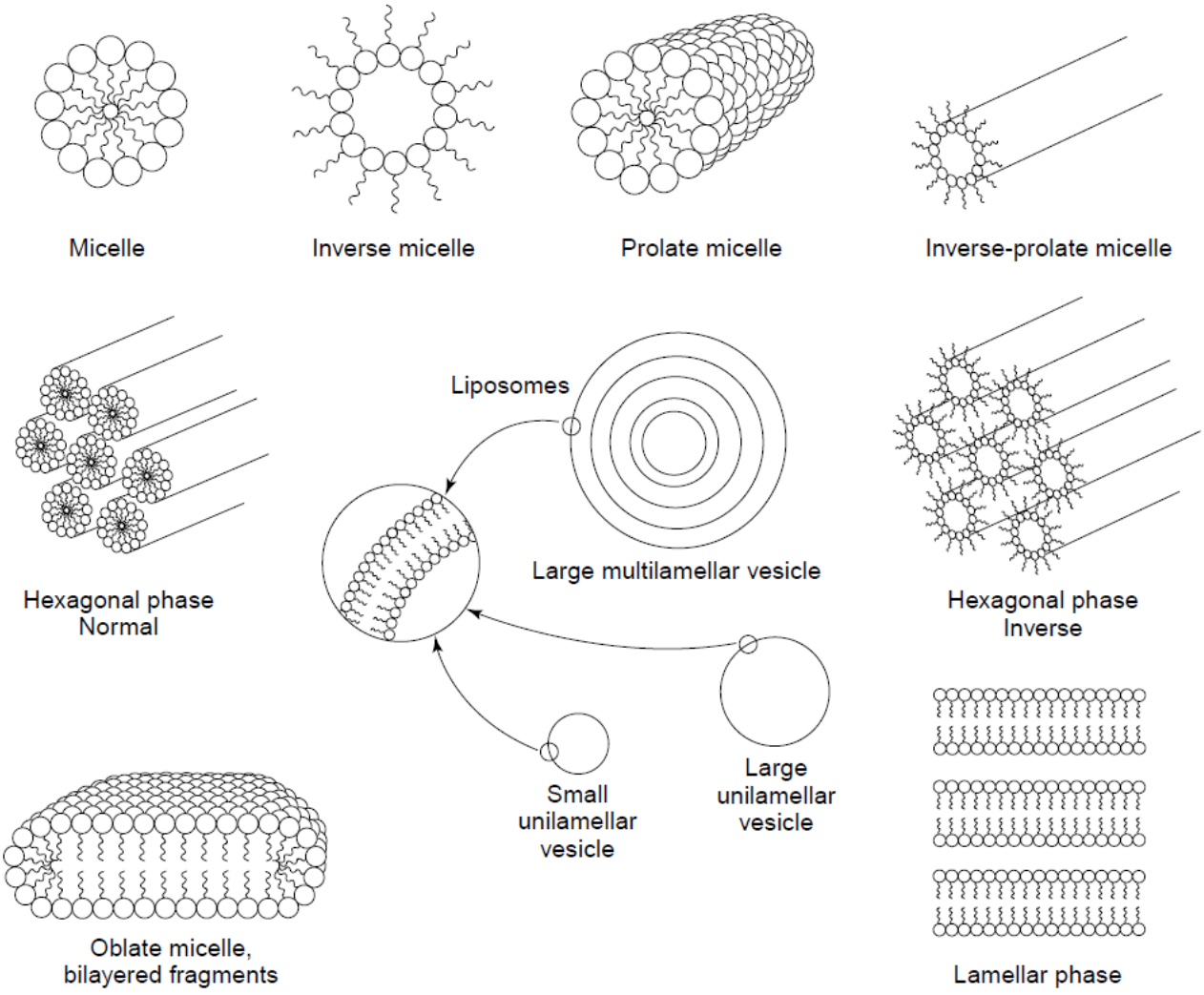

Figure 2.1: Examples of self-assembled aggregates and macroscopic liquid-crystalline phases formed by amphiphilic molecules in amphiphile-water mixtures. The self-assembly is strongly dependent on the molecular structure of the amphiphile, its concentration and the condition in the solution (temperature and ionic strenght). This illustration was adopted from reference 53.

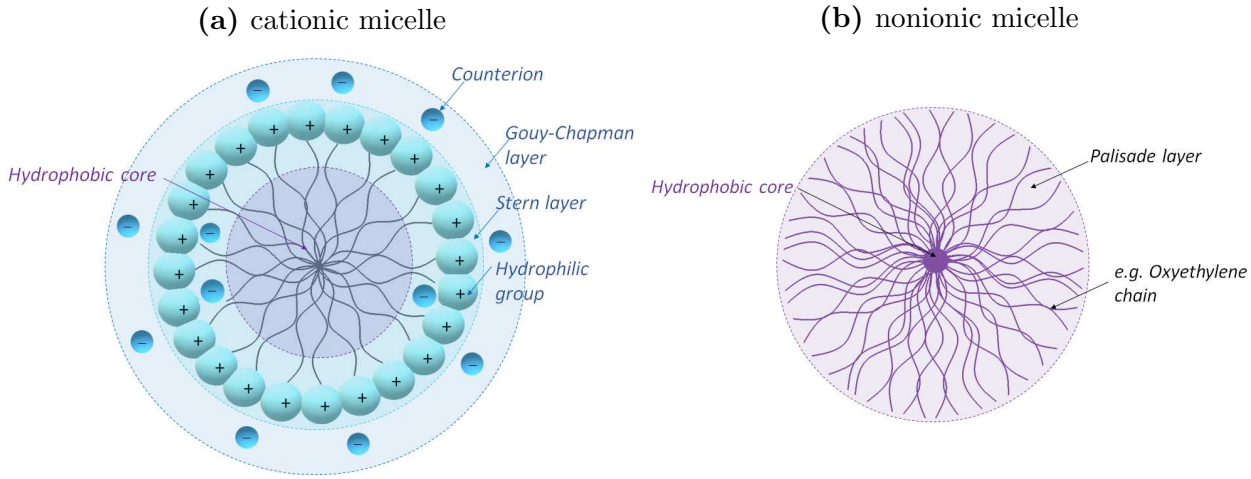

Figure 2.2: Schematic illustrations of cationic (a) and nonionic (b) micelle structures.

In nonionic micelles, the hydrophobic core is surrounded by the palisade layer, which is a shell containing the hydrated polar chains of the surfactant.[58] An important parameter that characterizes micelles is the aggregation number N_{agg}, which is the average number of surfactant monomers in the micelles of a micellar solution. The aggregation number, as well as the shape of the micelles can vary with the surfactant concentration.[59] With increasing the surfactant concentration, spherical micelles grow to various morphologies, most common are ellipsoidal (prolate), rod-like and worm-like micelles, bilayers fragments and reversed micelles.[60] Other factors affecting the micelle size and the CMC value are the structure of the hydrophobic and hydrophilic group, the type of the counterion in case of ionic micelles, the addition of electrolytes and the temperature of the solution.[58] An increase in the lenght of the hydrophobic chain results in an increase in the micelle size and a decrease in CMC. In general, nonionic surfactants have lower CMC values and higher aggregation numbers than ionic surfactants with similar hydrocarbon chains. An increase in the polar chain of nonionic surfactants leads to an increase in the CMC. For cationic surfactants, the micelle size increases depending on the type of the counterion according to the series $Cl^- < Br^- < I^-$ and for anionic surfactants according to the series $Na^+ < K^+ < Cs^+$.[58] The addition of electrolytes to ionic surfactant solutions leads to increasing micelle sizes and decreasing CMCs. In nonionic surfactant solutions at temperatures above the so-called cloud point temperature (CPT), a characteristic phase separation is observed, leading to two liquid phases: one is a diluted surfactant solution (micelle concentration around the CMC), the other is an aggregated, surfactant-rich phase.[61] Above the CPT, the micelle size is increased and the CMC is decreased.[58]

An important property of micelles with significance in pharmacy is their

ability to solubilize various sparingly water soluble substances by their incorporation into the micelle aggregates.[57] Due to the anisotropic micelle structure and water distribution within the micelle, the solutes can be located in different regions of the micelle depending on their chemical structure and polarity. Nonpolar molecules are solubilized in the hydrophobic core of ionic and nonionic micelles. Water insoluble compounds containing both polar and nonpolar groups are orientated with the polar group to the hydrophilic group of the micelle and the nonpolar group distributed inside the hydrophobic core. Solutes containing polar groups can be located in the outer shell of nonionic micelles, known as palisade layer (see Figure 2.2b).[57] Hydrophilic compounds can be adsorbed on the surface of the micelle as well as charged molecules, which are favorably distributed within the Stern layer of ionic micelles with the opposite charge (see Figure 2.2a).[57,60] The solubilization of a molecule i by a surfactant can be evaluated based on two descriptors: the molar solubilization capacity, χ, and the molar micelle/water partition coefficient, K_i, which are related to each other according to[57,62]

$$K_i = \frac{\chi(1 - \mathrm{CMC})}{S_\mathrm{w}} \tag{2.1}$$

where S_w is the water solubility of the solubilized compound. The molar solubilization capacity, χ, is defined as the number of moles of the solute that can be solubilized by one mole surfactant solution given with[57,63]

$$\chi = \frac{(S_\mathrm{tot} - S_\mathrm{w})}{(C_\mathrm{surf} - \mathrm{CMC})} \tag{2.2}$$

where S_tot is the total solute solubility and C_surf is the molar surfactant concentration. The solubilization capacity is influenced by factors such as the chemical structure of the surfactant and the solute, the surfactant concentration and the temperature of the solution.[57,58]

Liposomes are spherical vesicles, in which membranes consisting of one or more lipid bilayers encapsulate solvent molecules.[53] The major types of liposomes are the multilamellar vesicle (MLV), containing several lamellar lipid bilayers, small and large unilamellar vesicle (SUV and LUV, respectively) with one lipid bilayer (see Figure 2.1). In contrast to micelles, which form spontaneuosly by mixing surfactant and water and are thermodynamically stable systems, liposomes are not in thermodynamically stable state and therefore cannot form spontaneously. In order to produce liposomes some energy should be supplied to a dispersion of (phospho)lipids in a polar solvent (e.g., sonication, extrusion or homogenization).[53] Liposomes and lipid bilayers undergo several phase transitions. Above a certain temperature, known as phase transition temperature, T_c,

the lipid bilayers exist in a disordered liquid-crystalline phase, where the hydrocarbon chains are randomly orientated, whereas below T_c the ordered gel phase is observed, in which the hydrocarbon chains are fully extended and closely packed.[64,65] Moreover, different liquid-crystalline phases can be observed. By decreasing or increasing the lipid concentration, the lamellar phase can undergo a transition into a hexagonal micellar phase and an inverse hexagonal micellar phase, respectively.[66] These transitions can be induced by increasing or decreasing the average area of polar heads as well, for example by adjusting the pH value in case of pH-sensitive liposomes.[53,67,68] The above mentioned transitions result in liposome leakage of the encapsulated solvent, followed by their dissolution. As micelles, liposomes are able to solubilize solutes in aqueous solutions. In general, hydrophobic compounds are solubilized inside the lipid bilayers and hydrophilic compounds can be dissolved in the encapsulated aqueous content. Liposomes find various applications as delivery agents for encapsulated drugs due to the possibility to adjust their properties, such as leackage rate, fusogenic activity and interaction with cells for a particular application.[53]

2.2 Thermodynamics of micellar systems

2.2.1 Thermodynamics of micellization

The micelle formation is a reversible equilibrium process, characterized by a minimization of the free energy of the system in order to reach thermodynamic stability.[69] Experimental analysis showed only a small enthalpy change, when micellization occurs. Therefore, the process is considered as mainly entropy driven and the negative free energy results from a large positive entropy.[70] The micellization process is induced by hydrophobic, steric, electrostatic, hydrogen bonding, and van der Waals interactions.[57] The main attractive force results from an entropic effect, known as hydrophobic effect, which is the observed tendency of nonpolar molecules to aggregate in an aqueous solution and exclude water molecules, caused by the strong water self-association.[71,72] The process of micelle formation is characterized by the CMC, which is the surfactant specific concentration at which aggregates are spontaneously forming and become thermodynamically soluble in an aqueous solution.[73] The mass action model (MAM) and the pseudo phase separation approximation (PSA) postulate that the micelle formation can be modeled as a chemical equilibrium process between the micelles and the surfactant monomers. Hence, a relation between the CMC and the Gibbs free energy of micellization can be derived.[74] The key difference between the models is that the MAM provides the dependence of the chemical potential on the aggregation number N_{agg} and allows the consideration of

polydispersity.[75] The MAM gives more accurate description of the micellization and should be preffered for $N_{\mathrm{agg}} < 50$.[75] However, for large aggregation numbers ($N_{\mathrm{agg}} \to \infty$) or when the micelle becomes a macroscopic phase, the PSA can be interpreted as an approximation of the MAM. The Gibbs free energy of micellization for nonionic surfactants can then be defined as

$$\Delta G^0_{\mathrm{mic}} = \mu^0_{\mathrm{mic}} - \mu^0_{\mathrm{solv}} = RT \ln \mathrm{CMC} \tag{2.3}$$

At constant temperature and pressure, the chemical potential is equal to the partial molar Gibbs energy (see equation 2.18). Hence, μ^0_{mic} is equal to the free energy of inserting a surfactant into the micelle and μ^0_{solv} is equal to the solvation free energy of a surfactant in the aqueous phase (hydration free energy). Therefore, the CMC can be estimated from the free energy for the transfer of a surfactant from the aqueous phase to the micelle.[76] The formation of ionic micelles is strongly affected by the salt concentration. The ionic strenght will cause adsorption of some of the free counterions on the surface of the micelles. Therefore, the degree of counterion binding (β) is accounted for and the Gibbs free energy of micellization for ionic surfactants is approximated by[60]

$$\Delta G^0_{\mathrm{mic}} = (1 + \beta) RT \ln \mathrm{CMC} \tag{2.4}$$

Equation 2.4 can be further extended to account for the presence of electrolytes.[77]

Semiquantitative predictive models of micellization, such as the concept of the packing parameter by Israelachvili et al.[78], the Tanford theory[79] and its extention by Nagarajan and Ruckenstein[80] were introduced for the prediction of micelle shapes and sizes.

The free energy model by Tanford[79] suggest that the standard free energy change associated with the transfer of a surfactant from its infinite diluted state in water to an aggregate with aggregation number N_{agg} contains three contributions[81]

$$\left(\frac{\Delta\mu^0_{N_{\mathrm{agg}}}}{k_{\mathrm{B}}T}\right) = \left(\frac{\Delta\mu^0_{N_{\mathrm{agg}}}}{k_{\mathrm{B}}T}\right)_{\mathrm{Transfer}} + \left(\frac{\Delta\mu^0_{N_{\mathrm{agg}}}}{k_{\mathrm{B}}T}\right)_{\mathrm{Interface}} + \left(\frac{\Delta\mu^0_{N_{\mathrm{agg}}}}{k_{\mathrm{B}}T}\right)_{\mathrm{Head}} \tag{2.5}$$

The transfer term is a negative free energy contribution arising from the transfer of the surfactant tail from its unfavorable contact with water to the hydrocarbon environment in the aggregate core, which depends on the surfactant tail but not on the size and shape of the aggregate. The second term is a positive contribution, which takes into account that there is a residual contact of the tail surface area with water at the surface of the aggregate core. The

third term provides a positive contribution representing the repulsive interactions between the head groups at the aggregate surface due to steric interactions for all types of head groups and due to electrostatic interactions for zwitterionic and ionic head groups, as well. k_B is the Boltzmann constant, and T the temperature.

The packing paramater by Israelachvili et al.[78] is defined as

$$P = \frac{V_0}{a_e l_0} \tag{2.6}$$

whereby V_0 is the surfactant tail volume, a_e is the equilibrium head group area and l_0 is the extended tail lenght. a_e is a thermodynamic quantity obtained from equilibrium considerations (minimum free energy) and is not a simple geometrical parameter based on the shape and size of the surfactant head group.[81] Therefore, Israelachvili et al.[78] estimated a_e using the Tanford model for the standard free energy change on aggregation. For micelles in thermodynamic equilibrium, the equilibrium head group area a_e can then be calculated with[81]

$$a_e = \left(\frac{\alpha}{\sigma}\right)^{1/2} \tag{2.7}$$

where α is the head group repulsion parameter and σ is the interfacial tension, respectively. Based on the packing paramater value, the most probable geometries of the surfactant aggregates can be predicted: spherical or ellipsoidal micelles ($P < 0.33$), cylindrical or rod-like micelles ($P = 0.33 - 0.5$), vesicles and flexible bilayers ($P = 0.5 - 1.0$) and reversed micelles ($P > 1.0$). As for most common surfactants $V_0/l_0 =$ constant, only the head group area a_e controls the equilibrium aggregate structure and the surfactant tail does not have any influence on the shape and size of the aggregate in the model of Israelachvili et al.[78] The packing parameter concept combined with the Tanford's free energy expressions is able to describe various experimental findings. For example, for nonionic surfactants with ethylene oxide head groups, smaller head groups lead to small a_e and therefore to high P, which means bilayers are favored, where for larger head groups, P is lower and cylindrical micelles are favored.[81] In case of ionic surfactants, the addition of salt leads to a decrease in a_e and increase in P, therefore the transition from spherical to cylindrical micelles can be predicted.[81] The influence of solvents and temperature can be predicted, as well.[81] However, this concept will predict the same P value for surfactants with short and long tail lenghts and therefore still neglects the influence of the tail lenght on the shape and size of the aggregate.

Nagarajan and Ruckenstein[80] extended the Tanford model by introducing an additional term, the packing free energy contribution to the free energy

calculation. Taking into account this term, the estimated equilibrium head group area a_e is dependent on the lenght of the hydrophobic tail.[81] In this model, the tail length influences the head group area and thereby the shape of the aggregate. The extension by Nagarajan and Ruckenstein[80] improves the prediction quality of the semiquantitative predictive group of models, as the tail still has an influence on the formation of spherical micelles, rod-like micelles and bilayers. For example, the tail influences the ionic strenght in case of ionic micelles and thereby modifies the equilibrium area a_e, the packing parameter P and thus the aggregate size.[81]

Further extensively reported approaches for prediction of micellization (CMC) are quantitative structure-property relationship (QSPR)[82–86] models, activity coefficient models such as nonrandom two-liquid (NRTL)[87] and universal quasi-chemical functional-group activity coefficients (UNIFAC),[88] and equations of states (e.g., statistical associating fluid theory (SAFT)).[89,90]

2.2.2 Formation of micelles in mixed surfactant solutions

Mixtures of different surfactants are often used, as they can have improved characteristics compared to single surfactant solutions for a particular application.[29–31] One important issue in the formation of mixed micelles is their composition, which cannot be straightforward assumed as equal to the composition in the surfactant solution.[91,92] Different approaches have been proposed to calculate the CMC in mixed surfactant solutions and to predict the composition of the mixed micelles. If ideal mixing is assumed, the CMC value for a mixed surfactant system can be calculated with Clint's equation[93]

$$\frac{1}{\mathrm{CMC}_{12}} = \frac{\alpha_1}{\mathrm{CMC}_1} + \frac{(1-\alpha_1)}{\mathrm{CMC}_2} \tag{2.8}$$

where CMC_1 and CMC_2 are the CMCs of the pure surfactants 1 and 2, respectively and CMC_{12} is the CMC of the mixture. α_1 is the mole fraction of surfactant type 1 in the bulk solution. In case of ideal mixing, the mole fraction of surfactant 1 in the micelle x_1 in a solution with bulk composition α_1 is calculated as[93]

$$x_1 = \frac{\alpha_1 C_s - \mathrm{CMC}_1^*}{C_s - \mathrm{CMC}_{12}} \tag{2.9}$$

CMC_1^* indicates the CMC of surfactant 1 in a mixed solution with composition α_1, where $\mathrm{CMC}_{12} = \mathrm{CMC}_1^* + \mathrm{CMC}_2^*$.[94] C_s is the total surfactant concentration in the solution. The relation between x_1 and α_1 is given with

$$x_1 = \frac{\alpha_1 \mathrm{CMC}_2}{\alpha_1 \mathrm{CMC}_2 + (1-\alpha_1)\mathrm{CMC}_1} \tag{2.10}$$

The regular solution approximation theory (RSA) is another approach for the calculation of micelle compositions, which takes into account the non-ideality of mixing. The activity coefficients γ in a binary surfactant mixture can be describes according to the RSA by

$$\ln(\gamma_1) = \beta(1 - x_1)^2 \tag{2.11}$$

$$\ln(\gamma_2) = \beta(x_1)^2 \tag{2.12}$$

The parameter β, proposed by Rubingh[91] is a specific molecular interaction parameter, which accounts for the interaction energy difference between equal and nonequal surfactant molecules according to[95]

$$\beta = \frac{\ln\left[\frac{\alpha_1 \mathrm{CMC}_{12}}{x_1 \mathrm{CMC}_1}\right]}{(1-x_1)^2} = \frac{\ln\left[\frac{(1-\alpha_1)\mathrm{CMC}_{12}}{(1-x_1)\mathrm{CMC}_2}\right]}{x_1^2} \tag{2.13}$$

Based on the experimental data of the CMCs of the single surfactants and the mixture, as well as on the composition of the surfactant solution α_1, the composition x_1 of a micelle can be calculated numerically by

$$\frac{x_1^2 \ln((\mathrm{CMC}_{12}\alpha_1)/(\mathrm{CMC}_1 x_1))}{(1-x_1)^2 \ln((\mathrm{CMC}_{12}(1-\alpha_1))/(\mathrm{CMC}_2(1-x_1))} = 1 \tag{2.14}$$

Significant deviations between the composition in the bulk solution α_1 and the composition in the micelle x_i are obtained at total surfactant concentrations close to the CMC of the mixture ($C_\mathrm{s} \approx \mathrm{CMC}_{12}$), whereas at high concentrations ($C_\mathrm{s} \gg \mathrm{CMC}_{12}$) x_i approaches α_1.[93,96]

2.2.3 Partition equilibria in micellar systems

Phase equilibria in liquid systems and partition coefficients

In the state of thermodynamic equilibrium three conditions have to be fulfilled: the conditions of thermal, mechanical and chemical equlibrium.

$$T^\alpha = T^\beta = ... = T^\pi \tag{2.15}$$

$$p^\alpha = p^\beta = ... = p^\pi \tag{2.16}$$

$$\mu_i^\alpha = \mu_i^\beta = ... = \mu_i^\pi \tag{2.17}$$

The thermal and mechanical equilibria are characterized by the equality of the temperature T and pressure p of all phases according to equations 2.15 and

2.16, respectively. The chemical equilibrium is defined by the equality of the chemical potential of each component i in all phases, given with equation 2.17.

According to the Gibbs fundamental equations, the chemical potential of a component i can be defined based on the thermodynamic potentials, the internal energy $U(S, V, n_i)$, the enthalpy $H(S, p, n_i)$, the Helmholtz free energy $A(T, V, n_i)$ and the Gibbs free energy $G(T, p, n_i)$[97]

$$\mu_i = \left(\frac{\partial U}{\partial n_i}\right)_{S,V,n_j} = \left(\frac{\partial H}{\partial n_i}\right)_{S,p,n_j} = \left(\frac{\partial A}{\partial n_i}\right)_{T,V,n_j} = \left(\frac{\partial G}{\partial n_i}\right)_{T,p,n_j} \tag{2.18}$$

The chemical potential in an ideal mixture is calculated by

$$\mu_i^{\text{ideal}} = \mu_i^0 + RT \ln(x_i) \tag{2.19}$$

where μ_i^0 is the chemical potential of the chosen standard state, T is the temperature and x_i is the mole fraction of component i.

However, to describe real mixtures, the non-ideality has to be considered. Therefore, the activity a_i is introduced, which accounts for the non-ideality of the mixture. The chemical potential in a real mixture is then described as

$$\mu_i^{\text{real}} = \mu_i^0 + RT \ln(a_i) \tag{2.20}$$

By combining equations 2.17 and 2.20, the condition of chemical equilibrium can be expressed as

$$a_i^\alpha = a_i^\beta = ... = a_i^\pi \tag{2.21}$$

In order to quantify the non-ideality, the activity coefficient is defined as

$$\gamma_i = \frac{a_i}{x_i} \tag{2.22}$$

From equations 2.21 and 2.22, the general equation for the description of thermodynamic liquid-liquid phase equilibria (LLE) is derived

$$x_i^\alpha \gamma_i^\alpha = x_i^\beta \gamma_i^\beta = ... = x_i^\pi \gamma_i^\pi \tag{2.23}$$

The partition equilibria of a component i between two liquid phases α and β can be quantified with the partition coefficient (derived from equation 2.23)

$$K_{i,\text{sd}}^{\alpha\beta} = \frac{x_i^\alpha}{x_i^\beta} = \frac{\gamma_i^\beta}{\gamma_i^\alpha} \tag{2.24}$$

where the subscript "sd" stands for system size dependent. However, for a reasonable comparison, a system size independent partition coefficient is usually used, which is defined as the ratio of the molar concentrations c_i^α and c_i^β according to[98]

$$K_i^{\alpha\beta} = \frac{c_i^\alpha}{c_i^\beta} = \frac{x_i^\alpha}{x_i^\beta}\frac{v^\beta}{v^\alpha} = \frac{\gamma_i^\beta}{\gamma_i^\alpha}\frac{v^\beta}{v^\alpha} = K_{i,\mathrm{sd}}^{\alpha\beta}\frac{v^\beta}{v^\alpha} \tag{2.25}$$

Hence, the partition coefficients $K_{i,\mathrm{sd}}^{\alpha\beta}$ and $K_i^{\alpha\beta}$ are connected by the molar volumes v^α and v^β of the phases α and β, respectively. Throughout this work, the system size independent partition coefficient $K_i^{\alpha\beta}$ is used, where α represents the organic (micellar) phase and β is the aqueous phase, as established in the literature.[98]

Partition coefficients of ionizable solutes

Ionization reactions for acids AH can be expressed with

$$\mathrm{AH} \rightleftharpoons \mathrm{A^-} + \mathrm{H^+} \qquad\qquad K_\mathrm{a} = \frac{[\mathrm{A^-}][\mathrm{H^+}]}{[\mathrm{AH}]} \tag{2.26}$$

where the brackets represent the concentrations of each species and K_a the equilibrium constant of the acids. The association of ionic species is quantified analog to K_a with the ion pair formation constant K_IP

$$\mathrm{A^-} + \mathrm{X^+} \rightleftharpoons \mathrm{A^-X^+} \qquad\qquad K_\mathrm{IP} = \frac{[\mathrm{A^-X^+}]}{[\mathrm{A^-}][\mathrm{X^+}]} \tag{2.27}$$

where $[\mathrm{X^+}]$ and $[\mathrm{A^-X^+}]$ are the counter ion concentration and the ion pair complex concentration, respectively.

Combining equation 2.26 with the definitions of pH and pK_a results in the so-called Henderson-Hasselbach equation

$$\mathrm{pH} = \mathrm{p}K_\mathrm{a} + \log\Big(\frac{[\mathrm{A^-}]}{[\mathrm{AH}]}\Big) \tag{2.28}$$

As the ratio $[\mathrm{A^-}]/[\mathrm{AH}]$ is unitless, other ratios with other units may be used, for example the molar fractions of the dissociated and protonated form $x_{\mathrm{A^-}}$ and x_AH.

The partition coefficients of an unionized (protonated) compound $K_{i,\mathrm{AH}}^{\alpha\beta}$ and its ionized (dissociated) form $K_{i,\mathrm{A^-}}^{\alpha\beta}$ between two liquid phases α and β can be expressed with

$$K_{i,\mathrm{AH}}^{\alpha\beta} = \frac{[\mathrm{AH}]^\alpha}{[\mathrm{AH}]^\beta} \tag{2.29}$$

$$K_{i,\mathrm{A^-}}^{\alpha\beta} = \frac{[\mathrm{A^-}]^{\alpha}}{[\mathrm{A^-}]^{\beta}} \tag{2.30}$$

The partition coefficient of an ion pair is calculated analogically with

$$K_{i,\mathrm{IP}}^{\alpha\beta} = \frac{[\mathrm{A^-X^+}]^{\alpha}}{[\mathrm{A^-X^+}]^{\beta}} \tag{2.31}$$

The pH dependent partitioning of an ionizable compound in consideration of ion pairing can be described in terms of a distribution coefficients $D_i^{\alpha\beta}$ according to

$$D_i^{\alpha\beta} = \frac{[\mathrm{AH}]^{\alpha} + [\mathrm{A^-}]^{\alpha} + [\mathrm{A^-X^+}]^{\alpha}}{[\mathrm{AH}]^{\beta} + [\mathrm{A^-}]^{\beta} + [\mathrm{A^-X^+}]^{\beta}} \tag{2.32}$$

Combining the partition coefficients of the protonated, dissociated species and the ion pair complex (equations 2.29, 2.30 and 2.31) with equation 2.28, the pH dependent distribution coefficient in consideration of ion pairing can be expressed for acids as follows

$$D_{i,\mathrm{acid;IP}}^{\alpha\beta} = \frac{K_{i,\mathrm{AH}} 10^{(\mathrm{p}K_\mathrm{a} - \mathrm{pH})} + K_{i,\mathrm{IP}} 10^{(\log C_\mathrm{salt} - \mathrm{p}K_\mathrm{IP})} + K_{i,\mathrm{A^-}}}{1 + 10^{(\mathrm{p}K_\mathrm{a} - \mathrm{pH})} + 10^{(\log C_\mathrm{salt} - \mathrm{p}K_\mathrm{IP})}} \tag{2.33}$$

where C_salt is the salt concentration and $\mathrm{p}K_\mathrm{IP}$ is the pK value of the ion pair complex.

However, Ingram et al.[99] assumed that not only free ions $[\mathrm{A^-}]$, but also the associated ion pair $[\mathrm{A^-X^+}]$ sum up to the total ion concentration $[\mathrm{A^-}]$, implied in equation 2.26. In this case, the pH dependent distribution coefficient for acids is given with

$$D_{i,\mathrm{acid}}^{\alpha\beta} = \frac{K_{i,\mathrm{AH}} 10^{(\mathrm{p}K_\mathrm{a} - \mathrm{pH})} + K_{i,\mathrm{A^-}}}{1 + 10^{(\mathrm{p}K_\mathrm{a} - \mathrm{pH})}} \tag{2.34}$$

Of course, distribution coefficients for basic compounds can be derived analog. Diagrams, showing the distribution coefficients in logarithmic scale in dependence of pH are reffered to as lipophilicity profiles.[36]

Experimental determination of partition coefficients with micellar liquid chromatography

Micellar liquid chromatography (MLC) is a chromatographic separation technique that allows the determination of partition coefficients of solutes in micellar systems.[100] The separation in MLC is based on the interactions of the solutes with a stationary and a mobile phase. The partitioning of the solute molecules is determined by the three phase model introduced by Armstrong and Nome[101] which describes the distribution behavior of the solute between the stationary (S) phase, the micelles (M), and the aqueous bulk phase (W).[102] The retention in MLC depends on the partition equilibria of the solute between the three coexisting phases, where hydrophobic, electrostatic and dipole interactions of the solute with the micelles and the surfactant modified stationary phase influence its elution. It has been shown that in particular the interactions between ionic surfactants and charged solutes have a great impact on the retention times: very short retention times due to repulsive forces when the surfactant and the solute have the same charge and in case of opposite charges long retention times because the solute molecules strongly bind to the ionic surfactants due to electrostatic attraction forces.[103] Retention models, based on various properties such as surfactant concentration and elution volumes are used to describe the retention behavior of the solute and thus to determine quantitatively its partition behavior in micellar systems. The two most commonly used retention models are those proposed by Armstrong and Nome[101] and by Arunyanart and Love[104]. It has been demonstrated that both of them usually lead to comparable results.[41,105,106] In this work, the model by Armstrong and Nome[101] is used

$$\frac{V_\mathrm{s}}{V_\mathrm{e} - V_0} = \frac{v(K_i^\mathrm{MW} - 1)}{K_i^\mathrm{SW}} C_\mathrm{m} + \frac{1}{K_i^\mathrm{SW}} \tag{2.35}$$

where V_s, V_e and V_0 are the volume of the stationary phase, the solute's elution volume and the void volume, respectively. v and C_m are the surfactant molar volume and the micelle concentration, given with $C_\mathrm{m} = C_\mathrm{s} - \mathrm{CMC}$, where C_s is the total surfactant concentration in the mobile phase. Based on equation 2.35 the partition coefficients between the stationary phase and the aqueous bulk phase K_i^SW and between the micelles and the aqueous phase K_i^MW can be determined by plotting the volumes ratio againts varying micelle concentration. The partition coefficient K_i^SW is described by the reciprocal y-axis section b with

$$K_i^\mathrm{SW} = \frac{1}{b} \tag{2.36}$$

K_i^{MW} can be determined from the slope a of the interpolation line $f(x) = ax+b$ according to

$$K_i^{\mathrm{MW}} = \frac{aK_i^{\mathrm{SW}}}{v} + 1 \tag{2.37}$$

The partition coefficient between the stationary phase and the micelles K_i^{SM} results as a quotient between K_i^{SW} and K_i^{MW}

$$K_i^{\mathrm{SM}} = \frac{K_i^{\mathrm{SW}}}{K_i^{\mathrm{MW}}} \tag{2.38}$$

2.3 Molecular dynamics simulations

Statistical mechanics is a theoretical framework for prediction of macroscopic properties (pressure, temperature, phase equilibria, etc.) of a many body system from its microscopic behavior. When combined with the laws of thermodynamics, it is known as statistical thermodynamics or equilibrium statistical mechanics. Statistical thermodynamics connects the properties of specific microscopic configurations of a thermodynamic system (microstate) to its macroscopic properties (macrostate).[107] Hence, the macroscopic properties are averages over microstates. A macrostate is characterized by a probability distribution of all possible microstates within a certain statistical ensemble. This idea is the basic of the ensemble concept introduced by J. Willard Gibbs.[108] In other words, an ensemble is the collection of all possible mictostates sharing a set of defined macroscopic properties (e.g., the same number of particles, pressure and temperature). Therefore, macroscopic properties are calculated as averages over the systems in the ensemble. For example, some common thermodynamic ensembles are defined as[108]

1. **Microcanonical ensemble (NVE = constant)** The number of particles N, the system's volume V and the total energy in the system E are constant.

2. **Canonical ensemble (NVT = constant)** The number of particles N, the system's volume V and the temperature T are constant.

3. **Grand canonical ensemble (μVT = constant)** The chemical potential μ, the system's volume V and the temperature T are constant.

4. **Isothermal-isobaric ensemble (NPT = constant)** The number of particles N, the pressure P and the temperature T are constant.

An important quantity in statistical thermodynamics is the partition function, which describes the statistical properties of a system in thermodynamic equilibrium. The partition function represents the connection between macroscopic themodynamic properties and microscopic models in a particular statistical ensemble.[109] Hence, every thermodynamic property can be calculated by means of the partition function of the corresponding system. Unfortunately, in many cases it is not possible to obtain an analytical solution for the partition function. However, computer simulations can be used to solve this problem using nummerical methods.[110] Two commonly applied simulation techniques are: (i) Monte Carlo simulations, which generate ensemble members using the laws of statistical thermodynamics and macroscopic properties are calculated as ensemble averages and (ii) molecular dynamics simulations, which generate a phase space trajectory using the laws of classical mechanics and calculate macroscopic properties as time averages.[110]

Molecular dynamics (MD) simulations is a technique for determining equilibrium and transport properties of a classical many particle system. In case of all-atom MD simulations, the particles represent atoms. The classical equations of motion are solved nummerically in successive time intervals and therefore a time sequence of microstates is obtained. Solving the differential Newton's second law of motion

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \tag{2.39}$$

yields a chronologic progression of the microstates (trajectory), which describes how the positions and velocities of the atoms in the system change over time. The continuous trajectory of a N particles system moves in a $6N$-dimensional phase space ($3N$ coordinates r and $3N$ momenta p). Macroscopic properties of the system are calculated from the phase space trajectory as time averages.

To study the behavior of real systems with MD simulations, the interactions between the participating molecules have to be modeled. These are defined in a so-called force field, which is an essential part of each MD simulation. The force field, which contains analytical forms of the interactions and values for their parameters is a crucial factor for the prediction quality of the simulation. Therefore, force field parameters are usually adjusted such that simulations are able to reproduce physical properties of the system. Which properties (e.g., density, heat of evaporation, energy of solvation, partitioning free energy, etc.) are used as target values depends on the type of force field.[111–114]

2.3.1 Force fields

The force calculation is the most essential and time consuming part of every MD simulation. To describe the total energy of a system, the classical Hamiltonian $H(p, r)$ is used, which is the sum of the kinetic energy $K(p)$ and potential energy $V(r)$

$$H(p, r) = K(p) + V(r) \tag{2.40}$$

The kinetic energy term is independent of the particle position, if no further configurational restrictions are applied on the system[115] and it is calculated according to

$$K(p) = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} \tag{2.41}$$

The potential energy term $V(r)$ is described by the intermolecular and intramolecular interactions of the particles in the system and is dependent on the positions of the particles

$$V(r) = V_{\text{inter}}(r) + V_{\text{intra}}(r) \tag{2.42}$$

A schematic representation of the intermolecular interactions is shown in Figure 2.3.

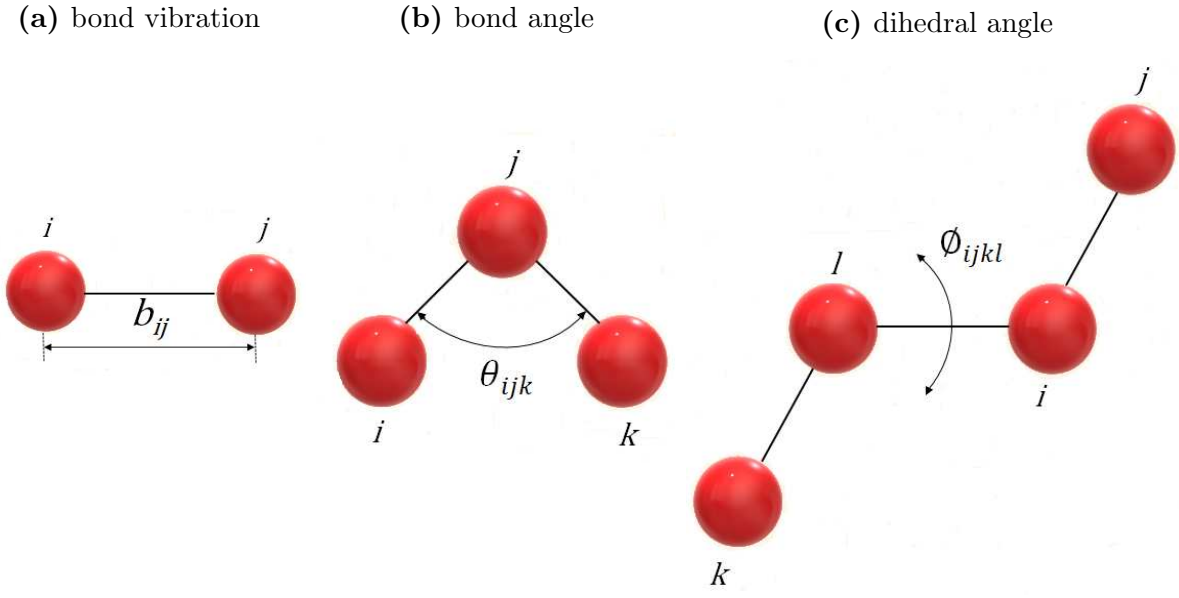

Figure 2.3: Schematic representation of the three most common intermolecular (bonded) interaction types.

The interactions between atoms and molecules and their parameter values are defined in the force field. Force fields for modeling biomolecules were proposed and established in the literature (e.g., GROMOS based force fields,[113,116–120] CHARMM,[112,121–123] AMBER,[124–126] OPLS[127,128]). In this work, CHARMM was chosen due to its good parametrization of amphiphilic molecules like lipids and surfactants.[121,129,130] The intra and inter molecular interactions in the potential energy function of the CHARMM force field[112] are calculated according to

$$V_{\text{intra}}(r) = \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{dihedrals}} K_\phi\left(1+\cos(n\phi-\delta)\right) + \sum_{\text{improper dihedrals}} K_\varphi(\varphi-\varphi_0)^2 + \sum_{\text{Urey-Bradley}} K_{\text{UB}}\left(r_{1,3}-r_{1,3;0}\right)^2 \tag{2.43}$$

$$V_{\text{inter}}(r) = \sum_{\text{nonbonded}} \left(\frac{q_i q_j}{4\pi D r_{ij}} + \epsilon_{ij}\left[\left(\frac{R_{\text{min},ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\text{min},ij}}{r_{ij}}\right)^{6} \right] \right) \tag{2.44}$$

The first term in the intramolecular energy function accounts for the bond stretches where K_b is the bond force constant, b_0 is the equilibrium distance and b the actual distance (see Figure 2.3a). The second term in equation 2.43 accounts for the bond angles where K_θ is the bond angle force constant, θ_0 is the equilibrium angle between three bonded atoms and θ is the bond angle. The third term is for the dihedral angles (torsion angles) where K_ϕ is the dihedral force constant, n is the multiplicity of the function, ϕ is the dihedral angle and δ is the phase shift. The fourth term accounts for the improper dihedral angles, which describe out of plane bending, where K_φ is the force constant, φ_0 is the equilibrium out of plane angle and φ the actual angle. The Urey-Bradley component (cross-term accounting for angle bending using 1,3 nonbonded interactions) comprises the fifth term, where K_{UB} is the respective force constant and r is the distance between the 1,3 atoms in the harmonic potential.

Nonbonded interactions between pairs of atoms (i, j) are represented by equation 2.44. The intermolecular interactions include the standard 12-6 Lennard-Jones (LJ) potential, where ϵ_{ij} is the well depth, $R_{\text{min},ij}$ is the radius in the LJ term used to treat the vdW interactions, and r_{ij} is the distance between i and j. The electrostatic energy is calculated with a Coulombic potential (see first term in equation 2.44), where q_i and q_j are the partial atomic charges of atom i and j, respectively. Since the number of nonbonded interactions increases quadratically with the number of simulated atoms, several techniques

to minimize the required computing power have been introduced.[110] The LJ potential is normally cut off at a certain distance, where a cutoff radius is defined and only the interactions with atoms within that defined sphere are taken into account.[110] Simply applying an abrupt truncation leads to discontinuity in the energy and in the forces. A possibility to avoid this problem is to shift the potential to zero at the cutoff radius.[110] Besides a shifting function a switching function can be used to truncate the van der Waals potential energy smoothly at the cutoff distance.[131,132] A switch distance specifies where the switching function should start taking effect to bring the van der Waals potential to zero smoothly at the cutoff distance. Thus, the switch distance must be less than the cutoff radius. Electrostatic interactions are treated by pairwise calculation with the Coulomb law. However, with increasing system size the calculation of all possible interactions would take most of the computational time and is not feasible. One possibility to overcome this problem is to truncate the forces at a cutoff distance, as used for the van der Waals potential calculation. The electrostatic interactions at distances greater than the cutoff can be accounted for using the particle mesh Ewald (PME) technique,[133] which calculates electrostatic potential beyond a chosen cutoff distance by the solution of a Poisson-Boltzmann equation for a solute in a continuum solvent. The electrostatic potential within the cutoff is calculated as Coulomb forces. Using PME allows a significant acceleration of the computing time.[134]

CHARMM General Force Field parametrization philosophy

The CHARMM force field was parametrized for varios classes of biomolecules including carbohydrates,[135–138] proteins,[122,139,140] nucleic acids[123,141,142] and lipids.[121,129] The main challenge in creating force fields is to cover enough of the variety of chemical classes. Therefore, the standard CHARMM optimization procedure for biomolecular force fields has been simplified for the purpose of creating the CHARMM General Force Field (CGenFF).[112] To attain this, a systematic optimization protocol is required that is (1) of a level of accuracy appropriate for the proposed application of the force field, (2) fast enough to parametrize large numbers of model compounds, and (3) simple enough to enable CHARMM users to easily extend the force field to chemical groups of their interest.[112] The CGenFF parametrization philosophy is adapted with emphasis on quantum mechanical (QM) results as target data for the parameter optimization.[112] This includes the QM calculations, followed by the analogous molecular mechanical (MM) calculations and parameter optimization to minimize the difference between the MM and QM data.[112] A general parametrization workflow for the force field parameter optimization according to the CGenFF procedure is shown in Figure 2.4 and the essential steps are summarized below.

1. **Generating initial topology and parameters**

   The parameterization workflow requires files containing molecular information such as initial molecular geometry, preassigned atom types and atom names.[143] For that, an automated algoritm based on the CGenFF is provided by the ParamChem[144,145] webserver, which is proven to be accurate for CGenFF atom types.[143] In many cases, it is not necessary to parametrize whole new molecules, but to undertake the parameter optimization procedure on only the new parameters required for the molecule of interest. That is, parameters missing in the force field. Initial values for the new parameters should be based on analogous parameters for similar chemical species already available in CGenFF. These can also be obtained with the ParamChem tool.[144,145] The initial molecular geometry should be optimized at the second order Møller-Plesset[146] level of theory on the 6-31G(d) basis set (further reffered as MP2/6-31G(d)).

2. **Optimization of partial atomic charges**

Optimization of the charges is performed with the model compound in its MP2/6-31G(d) optimized conformation. Complexes between the geometry-optimized target compound and a water molecule (TIP3P water model[147]) are constructed in which the water is ideally oriented for hydrogen bonding (distance and angle between acceptor and donor are optimized). For each target compound-water complex, a corresponding Hartree-Fock[148] calculation on the 6-31(d) basis set (further reffered as HF/6-31(d)) is performed. Scaled HF/6-31(d) model compound-water interactions and the dipole moment are used as target data. Ideally, the model compound-water interaction energies should be within 0.2 kcal/mol from the target interaction energies. For polar, neutral molecules, empirical results should overestimate the magnitude of the QM dipole moment by 20-50 % and should reproduce its orientation. Several rules facilitate charge fitting. For example, aliphatic hydrogen atoms are always assigned a charge of +0.09. Similarly, aromatic C-H moieties not adjacent to a heteroatom are assigned charges of -0.115 and +0.115 on the C and the H atom, respectively.

3. **Optimization of bonds and angles**

Optimized equilibrium bond and valence angle parameters should reproduce the target geometries, computed at the MP2/6-31G(d) level of theory. In general, deviations of up to 0.03 Å and 3° from the QM geometry are acceptable.

4. **Optimization of dihedrals**

Starting from the optimized MP2/6-31G(d) geometry, a potential energy scan (PES) for each dihedral angle has to be optimized at the MP2/6-31G(d) level of theory. The dihedral optimization is then performed until the MM PES reproduce the QM PES.

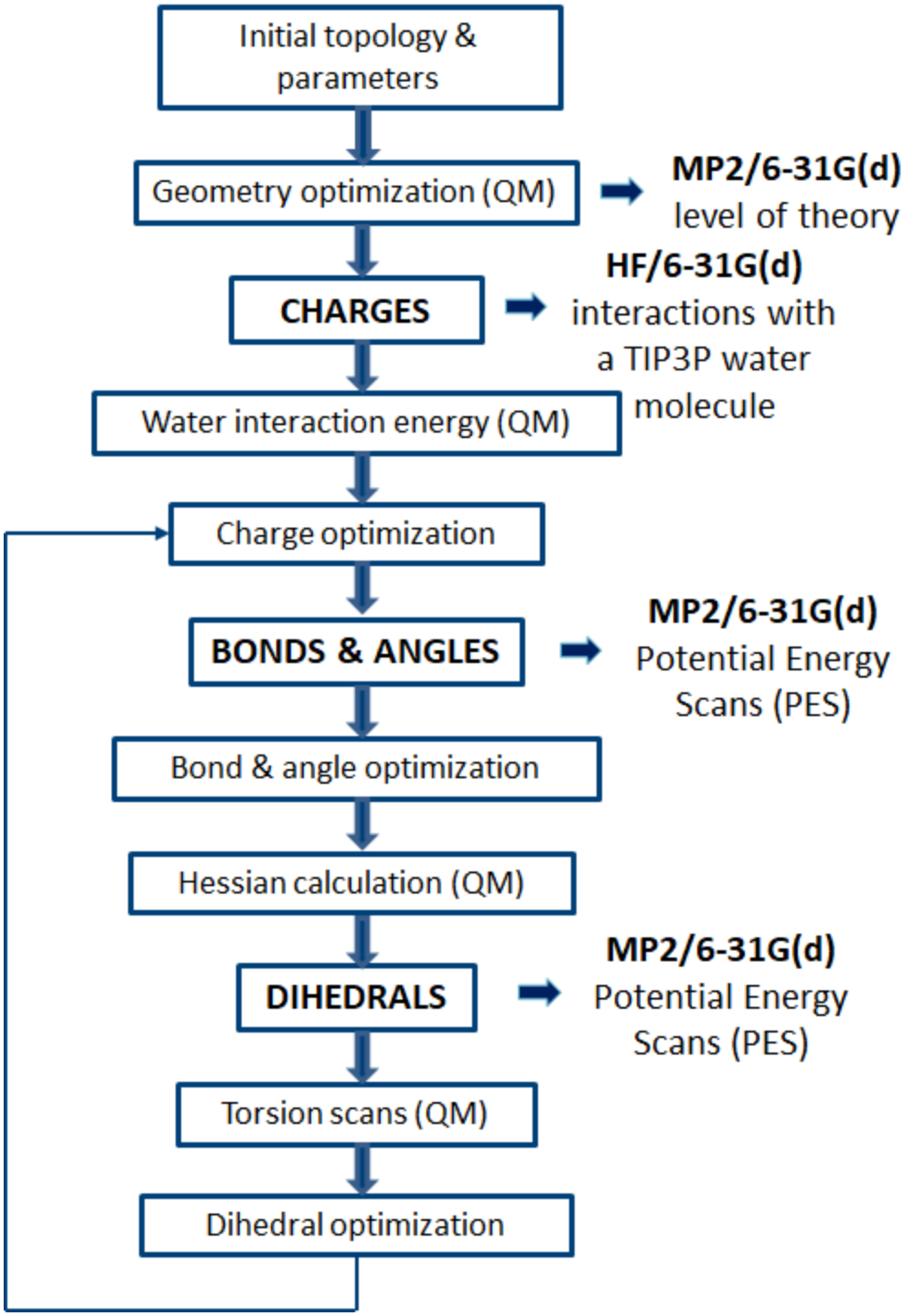

Figure 2.4: A general parametrization workflow according to the CGenFF procedure.[112,143]

In principle, all parameters in a force field are interdependent and need to be optimized in a self-consistent fashion. Therefore, force field parametrization is an iterative procedure (it is iterated from optimization of partial charges to optimization of dihedrals).[111] While this iterative procedure should be performed until convergence, with given good initial guess parameters, one or two iterations should be sufficient.[112]

2.3.2 Position restraints

In order to access areas with high energy barriers or to influence the overall structure of the simulated system, atoms or whole molecules can be restrained during a MD simulation. Atoms are restrained to fixed reference positions using a harmonic potential[149]

$$V_{pr}(r_i) = \frac{1}{2} K_{pr} |r_i - R_i|^2 \tag{2.45}$$

where r_i is the position of the atom, R_i is the reference position and K_{pr} is the force constant.

Simulations with position restraints find application during equilibration in order to avoid drastic rearrangements of critical parts.[149] Another important example of simulations with position restraints are the so-called pulling simulations, which are often carried out to generate initial configurations for umbrella sampling simulation windows (see the following chapter 2.3.3). In this case, a pulling force between two objects (molecules, clusters of molecules etc.) is applied, so that the relative position between them changes. In this work, a pulling force between the center of mass (COM) of a micelle and a solute molecule is applied, so that the solute approaches towards or away from the COM of the micelle.

Besides a harmonic potential, a flat-bottom (FB) potential can be used to prevent strong influences due to restrictions in the molecular movement. The FB potential consists of an inner part, in which the atom can move unhindered, and an outer part. If the atom leaves the inner region fixed with the radius r_{fb}, a harmonic potential is applied to return the atom to the inner region by[149]

$$V_{fb}(r_i) = \frac{1}{2} K_{fb} [d_g(r_i, R_i) - r_{fb}]^2 H[d_g(r_i, R_i) - r_{fb}] \tag{2.46}$$

where r_i is the position of the atom, R_i is the reference position, K_{fb} is the force constant and H the Heaviside step function. The distance $d_g(r_i, R_i)$ depends on the geometry of the FB potential. For a sphere it is calculated according to

$$d_g(r_i, R_i) = |r_i - R_i| \tag{2.47}$$

In this work, simulations with position restraints using a flat bottom potential are applied in chapter 6.2.2.

2.3.3 Free energy calculations

The Gibbs free energy G is a thermodynamic potential that is minimized when a system reaches equilibrium at constant pressure and temperature.[97] Hence, processes at constant pressure and temperature proceed spontaneously only, if G is reduced. The free energy along a reaction coordinate is known as a free energy profile and can be calculated from equilibrium MD simulations. A favorable solute partitioning can be estimated based on the probability of finding a solute at certain positions along the reaction coordinate. However, regions with high energy barriers (where solutes would not partition favorably) would require an extensively long simulation time to sample a solute in these regions. Different techniques to solve this problem were proposed in the literature (e.g., thermodynamic integration,[150] free energy perturbation,[151] umbrella sampling[152]). In this work, the umbrella sampling (US) method by Torrie and Valleau[152] is used.

Umbrella sampling (US)

The calculation of free energy profiles with US consists of several independent simulations (simulation windows). Each simulation window simulates the solute molecules at different positions along the reaction coordinate, therefore each window covers only a small range of the reaction coordinate. Thus, in each simulation window a different region of the reaction coordinate is sampled individually. Afterwards, the results from the independent simulations are combined to obtain a free energy profile (i.e., free energy along the reaction coordinate). In the US simulations, the reaction coordinate is restrained by applying a bias potential (in this work a harmonic potential). This additional energy term ensures efficient sampling along the whole reaction coordinate. The potential energy of the system with an applied bias potential $V^b(r)$ differs from the one of an unbiased system $V(r)$

$$V^b(r) = V(r) + w_i(\xi) \tag{2.48}$$

$w_i(\xi)$ is the bias potential, which depends only on the reaction coordinate.

When a harmonic potential is used, this additional term is calculated with

$$w_i(\xi) = \frac{1}{2} K \left(\xi - \xi_i^{\text{ref}}\right)^2 \tag{2.49}$$

where K is the force constant of the potential, ξ is the current position on the reaction coordinate and ξ_i^{ref} is the reference position in the respective window i. Since this harmonic potential only restricts the solute along the reaction coordinate, the molecule can continue to rotate and move freely in the orthogonal directions to the reaction coordinate.

In order to obtain the unbiased free energy, the unbiased distribution $P_i^u(\xi)$ is needed according to[153]

$$P_i^u(\xi) = \frac{\int \exp[-\beta V(r)]\delta[\xi\prime(r) - \xi]dN_r}{\int \exp[-\beta V(r)]dN_r} \tag{2.50}$$

where $\beta = 1/(k_\mathrm{B}T)$ with k_B being the Boltzmann constant and T being the absolute temperature. N_r is the number of degrees of freedom of the system.

The biased probability density $P_i^b(\xi)$ is calculated according to[153]

$$P_i^b(\xi) = \exp[-\beta w_i(\xi)]\frac{\int \exp[-\beta V(r)]\delta[\xi\prime(r) - \xi]dN_r}{\int \exp\big\{-\beta[V(r) + w_i(\xi\prime(r))]\big\}dN_r} \tag{2.51}$$

The unbiased probability $P_i^u(\xi)$ is related to the biased probability $P_i^b(\xi)$ according to[153]

$$P_i^u(\xi) = P_i^b(\xi)\exp[\beta w_i(\xi)]\langle\exp[-\beta w_i(\xi)]\rangle \tag{2.52}$$

The free energy $G_i(\xi)$ in simulation window i can then be determined via the biased probability density $P_i^b(\xi)$ with[153]

$$G_i(\xi) = -(1/\beta)\ln P_i^b(\xi) - w_i(\xi) + F_i \tag{2.53}$$

where $F_i = -(1/\beta)\ln\langle\exp[-\beta w_i(\xi)]\rangle$.

If the free energies $G_i(\xi)$ of more windows have to be combined to one global $G(\xi)$, F_i have to be calculated to connect $G_i(\xi)$ obtained in the different windows according to[153]

$$\begin{aligned}\exp(-\beta F_i) &= \langle\exp[-\beta w_i(\xi)]\rangle \\ &= \int P^u(\xi)\exp[-\beta w_i(\xi)]d\xi \\ &= \int \exp\big\{-\beta[G(\xi) + w_i(\xi)]\big\}d\xi\end{aligned} \tag{2.54}$$

where $P^u(\xi)$ is the global unbiased distribution (see equation 2.55).

As the F_i cannot directly be obtained from sampling, methods to calculate them (i.e., to analyze US simulations) have to be applied. These methods combine the results of different windows to reconstruct the entire system. Typically, the weighted histogram analysis method (WHAM)[154] or the umbrella integration[155] are used. In this work, the WHAM method is applied.

More details about the umbrella sampling technique can be found in the literature, for example in references 152, 153, 156, 157.

Weighted Histogram Analysis Method

WHAM[154] is a standard technique for generating free energy profiles from a set of US simulation windows. The global unbiased distribution $P^u(\xi)$ is calculated as a weighted average of the probabilities of the individual windows by[153]

$$P^u(\xi) = \sum_{i}^{\text{windows}} p_i(\xi) P_i^u(\xi) \tag{2.55}$$

p_i are the weights, chosen in order to minimize the statistical error of $P^u(\xi)$

$$\frac{\partial \sigma^2(P^u)}{\partial p_i} = 0 \tag{2.56}$$

under the condition $\sum p_i = 1$. They are calculated as[154,158]

$$p_i = \frac{a_i}{\sum_j a_j}, a_i(\xi) = N_i \exp[-\beta w_i(\xi) + \beta F_i] \tag{2.57}$$

with N_i being the total number of steps sampled for window i. The constants F_i (see also equation 2.53) are calculated by

$$\exp(-\beta F_i) = \int P^u(\xi) \exp[-\beta w_i(\xi)] d\xi \tag{2.58}$$

As there are two quantities depending on each other ($P^u(\xi)$ and the constants F_i), it has to be solved iterativelly. The global free energy of the system, calculated using WHAM is then defined as

$$G^{\text{WHAM}}(\xi) = -(1/\beta) \ln P^u(\xi) \tag{2.59}$$

For spherical structures, such as micelles an additional correction for the free energy is necessary. If the three-dimensional distance is devided in layers with a constant bin width for WHAM, their volume (volume of spherical shells) is not constant. Therefore, a correction is necessary to ensure that the free energy in the bulk water phase becomes constant. In this case, the effect of transforming Cartesian three-dimensional coordinates into a distance reaction coordinate ξ can be corrected with the Jacobian correction[159,160]

$$G(\xi) = G^{\text{WHAM}}(\xi) + (2/\beta) \ln \xi \tag{2.60}$$

More details about the volume correction for micelles are given in chapter 4.4.

Partition coefficients calculation

From free energy profiles the favorable partitioning of solutes can be quantified by means of partition coefficients. For their calculation, two phases have to be present: a micelle phase α and an aqueous phase. As free energy profiles are calculated along the whole reaction coordinate, a boundary between the two phases has to be defined. This can be specified based on the free energy values (a free energy of zero corresponds to the water phase). For the calculation of partition coefficients from free energy profiles in most MD studies the following equation is applied

$$K_i^{\mathrm{Pdef}} = \frac{V^{\mathrm{W}}}{m^{\mathrm{M}}} \frac{\sum_{j=1}^{n_\alpha} V(r_j) \exp(-\beta \Delta G_i(r_j))}{\sum_{k=n_\alpha+1}^{n} V(r_k) \exp(-\beta \Delta G_i(r_k))} \tag{2.61}$$

where the sum over layer 1 to n_α is defined to represent the micelle phase and the other layers ($n_\alpha + 1$ to n) define the aqueous phase. $V(r_j)$ and $V(r_k)$ denote the volume of shell j and k, respectively. V^{W} is the volume of the water phase and m^{M} denotes the mass of the micellar phase. K_i^{Pdef} is a system size independent partition coefficient based on the molality of the solute in the micelle phase and the molar concentration of the solute in the water phase in units L/kg.[161] The factor $V^{\mathrm{W}}/m^{\mathrm{M}}$ transforms the partition coefficient from a system size dependent into a system size independent coefficient.

2.4 COSMO-RS and COSMOmic

Besides MD simulations, the model COSMO-RS and its extention COSMOmic are used for the calculation of partition equilibria of solutes in this work. The models are based on quantum chemical calculations and statistical thermodynamics. Klamt[49] proposed the COSMO-RS model, which is developed on the basis of the original COSMO (COnductor-like Screening MOdel[162]). COSMO is a variant of the continuum solvation models, which rely on the assumption that the surrounding liquid environment can be approximated by a dielectric continuum, characterized by its dielectric constant ε_r. Quantum chemical calculations in a virtual conductor environment are performed as a first step to calculate the screening charge density on the surface of the molecule. In this environment, all interactions between the molecules are screened on the interface to the ideal conductor and there are no intermolecular interactions. However, in order to describe realistic solutions the intermolecular interactions have to be taken into account. Therefore, the COSMO theory was combined with the laws of statisti-

cal thermodynamics and extended to the COSMO-RS (COnductor-like Screening MOdel for Realistic Solvation)[49] approach. When using the COSMO-RS method for micellar systems, the pseudo phase approach[51] is applied, where the partition equilibrium of a solute molecule between two separate liquid phases (a surfactant and an aqueous phase) is calculated. In this case, the micelle is considered as a homogeneous macroscopic phase, whereas in reality micelles are anisotropic aggregates. Therefore, Klamt et al.[52] developed an extension of the COSMO-RS approach called COSMOmic, which takes into account the anisotropy of the system. In the next section, the COSMO-RS methodology is summarized briefly, whereas a detailed description can be found in references 49,163–167. Thereafter, in section 2.4.2, the COSMOmic approach is described.

2.4.1 The Conductor-like Screening Model for Realistic Solvation

In the COSMO-RS approach, the conductor environment ($\varepsilon_r = \infty$) is considered as a reference state for solvation. The screening charge density on the surface of a molecule in an ideal conductor is calculated using the density functional theory (DFT). The three-dimensional cavity around the molecule is then discretized into surface segments with a screening charge density σ for each segment. The sum of all segment charge densities σ gives a σ-profile (probability p(σ)), which provides information on the charge density distribution of the specific molecule. As an example, the σ-profile of water is shown in Figure 2.5. The two high peaks indicate similar positive and negative charge densities, which corresponds to the strong polarity of water.

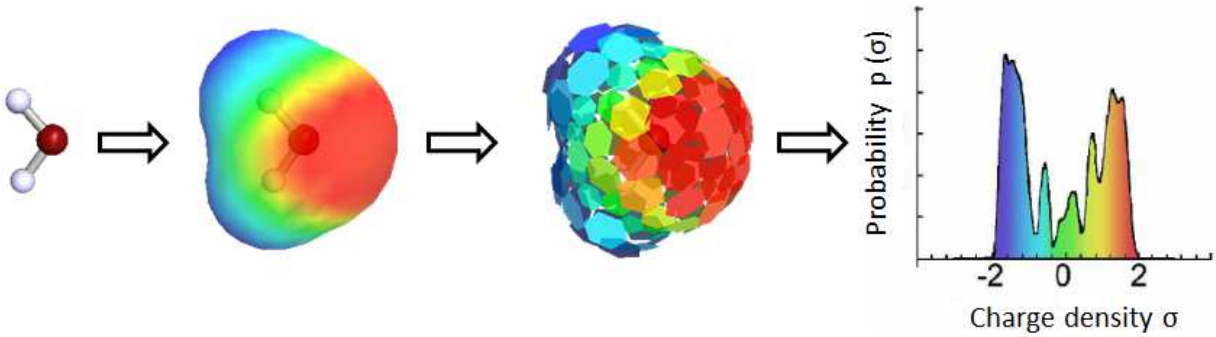

Figure 2.5: A schematic representation of the COSMO-RS charge density distribution calculation for the water molecule.[168]

The σ-profile provides the basis for the COSMO-RS calculations, where the short range interactions between different molecules are quantified by the interation energies of pairwise interacting segments. Several contributions to the interation energy are considered: an electrostatic contribution e_{misfit}, a hydrogen bonding contribution e_{hb} and a van der Waals contribution e_{vdW}.

The electrostatic energy contribution is decribed by the misfit energy term e_{misfit}. The misfit energy term is necessary, because it accounts for the energy difference between the ideal screening of the molecule in the ideal conductor and the non-ideal screening in a real solution. It is calculated according to

$$e_{\text{misfit}}(\sigma, \sigma') = \frac{\alpha'}{2}(\sigma + \sigma')^2 \tag{2.62}$$

where σ and σ' are the screening charge densities of two contacting segments, scaled with the misfit energy parameter α'. The misfit energy parameter is an empirical factor, fitted to experimental data.

For the hydrogen bonding contribution, an empirical term e_{hb} was proposed, given with equation 2.63. This term needs to be considered, because the interaction energy contribution due to hydrogen bonding cannot be neglected for interactions in liquid phases. The screening charge densities of both hydrogen bond acceptor σ_{acc} and donor σ_{don} should exceed a certain threshold value σ_{hb}. Otherwise, the e_{hb} term is equal to zero.

$$e_{\text{hb}}(\sigma, \sigma') = c_{\text{hb}}(T) \min(0; \sigma_{\text{don}} + \sigma_{\text{hb}}) \max(0; \sigma_{\text{acc}} - \sigma_{\text{hb}}) \tag{2.63}$$

where the hydrogen bonding energy factor c_{hb} is an empirical correction, which accounts for the temperature dependence.

The third energy contribution due to van der Waals interactions, caused by induced dipoles cannot be described based on the screening charge density of the surface segments. This interaction energy contribution is defined as

$$e_{\text{vdW}}(e, e') = \tau_{\text{vdW}}(e) + \tau_{\text{vdW}}(e') \tag{2.64}$$

whereas the van der Waals parameter τ_{vdW} has been fitted to experimental element specific parameters.

The sum of the three described interaction energy contributions gives the overall interation energy between two contacting surface segments

$$e_{\text{int}}(\sigma, \sigma') = e_{\text{misfit}}(\sigma, \sigma') + e_{\text{hb}}(\sigma, \sigma') + e_{\text{vdW}}(e, e') \tag{2.65}$$

Applying the principles of statistical thermodynamics, the surface specific chemical potential $\mu'(\sigma)$ of a segment σ is derived from the interactions with all segments in an ensemble.

$$\mu'(\sigma) = -\frac{k_{\text{B}}T}{a_{\text{eff}}} \ln\left(\int p'_{\text{s}}(\sigma') \exp\left(\frac{-a_{\text{eff}}(e_{\text{int}}(\sigma, \sigma') - \mu'(\sigma'))}{k_{\text{B}}T} \right) d\sigma' \right) \tag{2.66}$$

where $p'_{\text{s}}(\sigma')$ is the normalized σ-profile of the system, k_{B} and a_{eff} are the Boltzmann constant and the effective contact area, respectively.

The chemical potential of component i in a system is calculated as

$$\mu_i = \mu_{i,\text{res}} + \mu_{i,\text{comb}} + k_\text{B}T\ln(x_i) \tag{2.67}$$

The residual part of the chemical potential $\mu_{i,\text{res}}$ can be calculated as a sum of the surface segments chemical potentials according to

$$\mu_{i,\text{res}} = \int p_i(\sigma)\mu'(\sigma)d\sigma \tag{2.68}$$

where $p_i(\sigma)$ is the σ-profile of component i. The combinatorial contribution $\mu_{i,\text{comb}}$ is described by the Staverman-Guggenheim expression and accounts for the size and shape effects of a component in a liquid mixture.[169] x_i is the mole fraction of component i in the mixture.

Finally, the activity coefficient γ_i of component i is calculated by

$$\gamma_i = \exp\left(\frac{\mu_i - \mu_i^0 - k_\text{B}T\ln(x_i)}{k_\text{B}T}\right) \tag{2.69}$$

where μ_i^0 is the chemical potential of the standard state.

Partition coefficients between two liquid phases can then be derived from the activity coefficients (see equation 2.25).

2.4.2 COSMOmic for anisotropic systems

COSMOmic is an extension of COSMO-RS for anisotropic systems, such as micelles and membranes. A three-dimensional structure of the micelle or membrane is needed for the calculation with COSMOmic, which is usually obtained from a MD simulation (see chapter 4.3). For micelles, the three-dimensional structure is radially divided into several shells (layers). In case of membranes, the lipid bilayer structure is divided into layers along the membrane normal. The COSMOmic division of the input structures into layers is shown schematic in Figure 2.6. The structure of each layer is assumed to be homogeneous with a certain atomic composition of surfactant or lipid, water and ions (in case ions are present). COSMO-RS calculations of the solute in different orientations are performed for each layer, whereby solutes can expand over multiple layers. The solute molecules are considered in an infinite dilution, so that they cannot interact with each other.

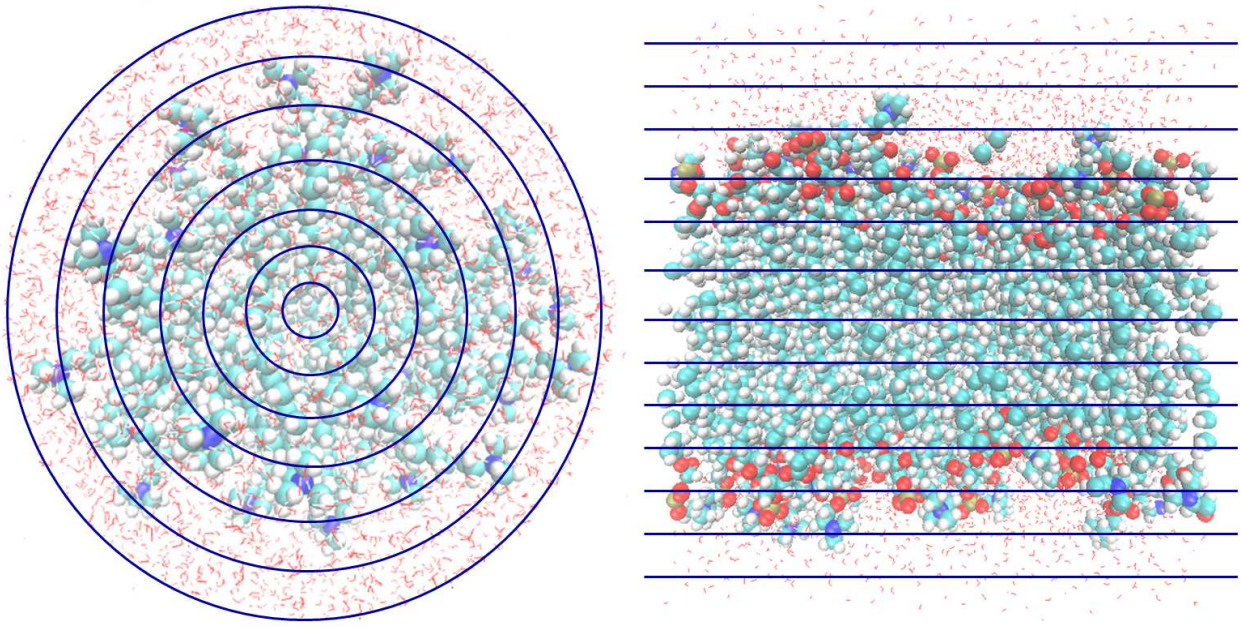

Figure 2.6: A schematic representation of the COSMOmic division into layers for micelles (left) and for membranes (right).

The chemical potential of solute i at a defined position r and orientation d can be calculated with

$$\mu_i(r,d) = \mu_{i,\mathrm{CRS}}(r,d) + \mu_{i,\mathrm{comb}}(r,d) \tag{2.70}$$

$\mu_{i,\mathrm{CRS}}$ represents the term of the chemical potential, which results from the pairing of surface segments and is often the dominating part of the chemical potential, escpecially for polar components.[52,170] The term $\mu_{i,\mathrm{comb}}$ takes into account the size and shape of the solute and solvent molecules. Originally, Klamt et al.[52] introduced two additional terms for the chemical potential calculation: an empirical approximate expression for the elastic deformation of the micelle or membrane by the penetration of a solute molecule and a zeta potential, taking into account the long range electrostatic interactions. Their influence was investigated and considered as negligible for neutral solutes.[52] However, for ionic species, the zeta potential should not be neglected and the calculation of charged solutes with COSMOmic should be interpreted with caution.[52] Bittermann et al.[171] introduced an approach to improve the COSMOmic prediction quality for charged molecules in membranes by implementing an internal membrane dipole potential (see also section 3.5).

To describe the partition behavior of a solute i, its chemical potentials in each layer j and each orientation k are summed up to the partition function

$$Z_i = \sum_{j=1}^{n} Z_i(r_j) = \sum_{j=1}^{n} \sum_{k=1}^{m} \exp\left(\frac{-\mu_i(r_j, d_k)}{k_\mathrm{B}T}\right) \tag{2.71}$$

The probability that the solute molecule partitions favorably in layer j can be determined by means of the distribution function

$$p_i(r_j) = \frac{Z_i(r_j)}{Z_i} \tag{2.72}$$

The free energy difference ΔG_i of the solute between the position r_j within the micelle or membrane and the aqueous bulk phase can be calculated from this probability with

$$\Delta G_i(r_j) = -k_\mathrm{B}T \ln p_i(r_j) \tag{2.73}$$

Combining the free energy differences at all position leads to the free energy profile. As for micelles the volumes of the layers are not constant, the volume correction is applied in this work to ensure that the free energy in the water phase is constant (see also chapter 4.4). In case of free energy profiles in membranes, no volume correction is needed, as the volume of the layers is constant.

For the calculation of partition coefficients with COSMOmic, Klamt et al.[52] introduced the concept of a water correction to obtain an equation that is stable also for hydrophilic solutes. In this case it is assumed that a fraction of the free energy in each layer is due to the water in this layer, and this fraction is subtracted leading to

$$K_i^{\mathrm{Wcorr}} = \frac{V^\mathrm{W}}{m^\mathrm{M}} \frac{\sum_{j=1}^{n} \left(V(r_j)\exp(-\beta\Delta G_i(r_j)) - V(r_n)\exp(-\beta\Delta G_i(r_n)) \left(\frac{n_\mathrm{w}(r_j)}{n_\mathrm{w}(r_n)}\right)\right)}{V(r_n)\exp(-\beta\Delta G_i(r_n))\frac{n_\mathrm{g}^\mathrm{w}}{n_\mathrm{w}(r_n)}} \tag{2.74}$$

whereby $V(r_j)$ denotes the volume of layer j, $n_\mathrm{w}(r_j)$ is the amount of water molecules in layer j, and n_g^w represents the total amount of water in the system. β is the reciprocal product of the Boltzmann constant and the temperature. Please note that in the COSMOtherm implementation of COSMOmic, the volume contribution is included in the calculation of the partition function with equation 2.71 and is not considered in equation 2.74. In equation 2.74 volume corrected free energy profiles should be used (see chapter 4.4). In case of calculating partition coefficients from unmodified profiles (profiles obtained with COSMOtherm), the volume contribution is already included in the partition function calculation and therefore in the free energy profile and should

not be considered in equation 2.74. K_i^{Wcorr} is a system size independent partition coefficient in units L/kg. The factor $V^{\mathrm{W}}/m^{\mathrm{M}}$, which is the volume of the water phase divided by the mass of the micellar phase transforms the partition coefficient from a system size dependent into a system size independent.

3 State of the art

3.1 Molecular dynamics simulations of micellization

Many MD simulations studies investigating the micellization process of various nonionic, zwitterionic, anionic and cationic surfactants were reported in the literature. De Nicola et al.[172] proposed and validated a coarse-grained model to estimate the CMC, as well as shape and phase transitions of the widely used nonionic surfactant TX100. Moreover, an atomistic model using the CHARMM force fields parameters optimized in this work (see chapter 5.1) and published in reference 173 was applied to calculate distance distribution functions and hydrodynamic measurements.[172] Polydisperse mixtures of spherical, oblate and prolate micelles were found, which is in agreement with experimentally observed micelle structures.[174–176] Espinosa et al.[177] employed atomistic MD simulations using the GROMOS force field parameters[178] to investigate the pressure effect on TX100 self-assembly process. At atmospheric pressure, a prolate ellipsoid micelle was obtained, whereas at higher pressures decreasing eccentricity of the micelle was calculated indicating more spherical structure. Changes of the micelle radius were observed with increasing pressure, which is related to the increase in the solvation of the hydrophilic area. Under high pressures the water density is modified, which makes it acting as an unstructured liquid capable of penetrating into the micelle.

Marrink et al.[179] performed MD simulations of dodecylphosphocholine (DPC) zwitterionic surfactant monomers at two concentrations above the CMC and obtained different aggregate shapes - a spherical micelle at the lower concentration and a worm-like micelle at the higher concentration.

Bruce et al.[180] evaluated structural characteristics (radius of gyration, eccentricity, micelle size, accessible surface area) as well as counterion distribution of a SDS micelle obtained from a 5 nanosecond all-atom MD simulation. Sammalkorpi et al.[181] studied the structural properties of SDS aggregates with large-scale MD simulations. A procedure to identify micelles from self-assembly simulations in aqueous solutions was proposed based on distances between the COMs of the surfactant monomers, as well as between chosen carbon atoms. Furthermore, the temperature effect on the self-aggregation process was investigated. Crystalline aggregates were found at low temperature and micelles were formed at elevated temperatures, respectively. Sammalkorpi et al.[182] studied the micellization of SDS in presence of salts (NaCl and $CaCl_2$) and found the the ionic strenght of the solution influences both the aggregation size and structure of the anionic micelles. Sanders et al.[183] investigated the micellization

behavior of sodium alkyl sulfates with different alkyl tail lenghts in water using atomistic self-assembly MD simulations. CMC values and aggregate size distribution were obtained at various temperatures. They found a strong dependence of the free monomer surfactant concentration on the overall surfactant concentration, which is crucial when comparing simulated and experimental CMCs. Espinosa Silva and Grigera[184] studied the formation of SDS aggregates at different pressures and temperatures. Sphere-like micelles were obtained from self-assembly simulations at atmospheric pressure and a temperature of 300 K. Larger rod-like shaped aggregates, which are pieces of bilayers were observed at high pressures. Both low and high temperatures lead to formation of lammelar structures instead of micelles. It could be demonstrated that low temperature has similar effect as high pressure, as expected due to the decreasing contribution of the entropic term. At very high temperature (700 K) formation of reverse micelles was observed, as water is in the vapor phase at this temperature.

The self-assembly process of SDS and CTAB surfactants at two different concentrations was investigated with all-atom MD simulations by Storm et al.[185] The expected concentration influence could be obtained (larger aggregates at higher concentrations). The reported aggregation number of both SDS and CTAB micelles are in the range of experimental findings and previous literature data.[181,186–188]

Storm et al.[170] studied the self-assembly of nonionic/ionic mixed micellar systems on an atomistic scale. They performed MD simulations of pure Brij35 and mixtures with CTAB and SDS at different concentrations and surfactant ratios. The obtained micelle sizes of Brij35 are in good agreement with experimental results. Moreover, the estimated composition of the mixed micelles is similar to the bulk surfactant composition, as expected.[94]

Chen and Hao[189] simulated CTAB and sodium octyl sulfate (SOS) mixtures with fixed number of surfactant molecules and different cationic/anionic surfactant ratios. The relation between the micelle composition and the shape of the aggregate was established, which contributes to the understanding of the self-assembly of mixed cationic/anionic surfactant systems.

MD simulations were used to investigate the self-assembly and the bilayer to micelle transition of lauric acid.[190] Formation of gel phase bilayer was observed at low and intermediate pH, whereas spherical micelles were obtained at high pH, which corresponds to the expectations that fatty acids form liposomes near their pK_a value and micelles, when the anionic form dominates, respectively.[191–193]

3.2 Free energy profiles in micelles

Precise information about the partition behavior in anisotropic environments can be obtained from free energy profiles, which are not accessible via experiments. Free energy profiles can give insights into the localization of solutes within the micelles depending on the solute-micelle combination. For lipid bilayers, MD simulations have been frequently used to calculate free energy profiles for many different molecules.[194] However, few publications applied MD simulations for free energy profiles in micellar aggregates so far. Matubayasi et al.[195] combined the method of energy representation with MD simulations to calculate the solvation free energy of methane, benzene and ethylbenzene in SDS micelles. The free energy of solvation shows that the solutes are energetically stabilized in the hydrophobic core regions. Fujimoto et al.[196] used MD simulations with thermodynamic integration and showed that methane is stabilized in the hydrophobic core of SDS micelles. Yan et al.[197] investigated the partition behavior of pyrene in SDS micelles using all-atom MD simulations. Free energy profiles showed that one pyrene molecule would partition in the hydrophobic core of the micelle, while two molecules build a conjugate structure (excited dimer) and are located near the head group region of the micelle. Yuan and Larson[198] studied flower-like micelles of a hydrophobically modified ethylene oxide urethane molecule ($C_{16}E_{45}C_{16}$) and star-like micelles containing $C_{16}E_{22}$ and $C_{16}E_{23}$ by atomistic and coarse-grained MD simulations. They used the US method to obtain free energy profile for transferring a hydrophobic group from the micelle to bulk water and estimated most probable micelle size and average hydrophobe escape time, consistent with experimental studies. Grafmüller et al.[199] performed atomistic MD simulations with the US method to obtain free energy profiles for the desorption of DPPC lipids from bilayers, as well as from a spherical micelle. It could be demonstrated that the desorption free energy strongly depends on the geometry of the lipid aggregate.

3.3 Partition equilibria prediction of neutral solutes in micelles

There are few approaches in the literature, which allow the prediction of micelle/water partition coefficients. One group are models based on linear correlations between physicochemical properties and partition coefficients (e.g., property-property relations (PPR)[45], linear solvation energy relations (LSERs)[46], quantitative structure-activity relationships (QSARs)[47,48]). However, these models require physicochemical properties of the molecules in the system.

Another group are models based on the pseudo phase approach (micelles are treated as a macroscopic phase in equilibrium with the aqueous surrounding). The partitioning of a solute between these coexisting phases is determined by the thermodynamic equilibrium, whereby activity coefficients can be calculated using g^E models (e.g., Universal Quasi-Chemical Functional-Group Activity Coefficient (UNIFAC)[88] or with the COSMO-RS model.[49,50,164–167] The COSMO-RS model has the advantage that only the chemical structures of the molecules have to be known for the calculation. Buggert et al.[200] studied the potential of UNIFAC and COSMO-RS to predict the partitioning of nonpolar solutes (toluene, p-xylene) in aqueous solutions of nonionic surfactants (TX100, Lutensol FSA10). The original UNIFAC model had been extended by an interfacial contribution term to take into account the micelle size and improved prediction quality could be demonstrated. The COSMO-RS model, which does not take into account the size of the micelles underestimates the concentration of the solutes in the aqueous phase. Mokrushina et al.[201] applied UNIFAC and COSMO-RS to predict micelle/water partition coefficients of homologous series of organic solutes in nonionic polyethoxy alcohols and SDS surfactant solutions. Good quantitative agreement with experimental data could be demonstrated. Moreover, factors that influence the prediction quality of COSMO-RS, such as the conformer of the molecules were discussed. In reference 202 two methods for performing conformational search for COSMO-RS were evaluated - the HyperChem program, which examines the molecules geometry in vacuum and MD simulations in solvent. MD simulations were considered as a promising tool for conformational analysis, as they account for the solvent in realistic manner. A methology to select conformers of large flexible molecules (surfactants, drugs, macromolecules) from MD simulations was developed and reproducible COSMO-RS predictions were achieved.[203] Good prediction quality of COSMO-RS was reported for partition coefficients of neutral solutes in Brij35 and CTAB micelles.[105] The method was evaluated for partition equilibria prediction in mixed Brij35/SDS and Brij35/CTAB micelles as well.[41,42] In case of mixed micelles, the composition of the mixed micelle has to be defined and it is a crucial factor for reasonable predictions. Wille et al.[204] applied COSMO-RS to predict the partitioning of alkyl parabens and ibuprofen in nonionic and ionic surfactant solutions with added organic and inorganic electrolytes. The experimentally obtained influence of added electrolytes on the partition behavior could be reproduced.

However, COSMO-RS employs the pseudo phase approach and therefore considers the micelle as a homogeneous surfactant phase, which is not the case in reality. The extension of the model called COSMOmic takes into account the anisotropy of the micelle based on the three-dimensional micelle structure

required for the calculation. Partition coefficients of neutral solutes in SDS and CTAB micelles as well as in mixed Brij35/SDS and Brij35/CTAB micelles were predicted with COSMOmic.[170,185,205] Very good agreement with experimental data and improved prediction quality compared to the pseudo phase approach were obtained.[205] Still, the micelle structure used for the COSMOmic calculation has a great impact on the prediction quality. Since the structures are taken from self-assembly MD simulations, different micelles occur over the simulation time and are available as an input for COSMOmic. It was demonstrated that averaged partition coefficients from calculations in different micelles with the same aggregation number obtained over the trajectory lead to more reproducible results than the partition coefficients calculated in single micelle structures.[185] In self-assembly simulations micelles offen are in close proximity to each other, which may lead to gaps in the water layer around the micelle and using such structures alters the predictions with COSMOmic.[170] The size and the shape of the micelle have high impact on the calculated values as well. Ritter et al.[206] analyzed the influence of micelle size and shape on the partition equilibria predictions in various micellar systems, where preassembled micelles with aggregation numbers between 20 and 160 for six different surfactants were considered. It was recommended to use small spherical micelles for the COSMOmic calculations, as outliers are more likely to occur when using larger and in particular more ellipsoidal aggregates.

3.4 Micelles composition of mixed surfactant systems

Thermodynamic theories for surfactant mixtures have focused on the CMCs of mixtures as well as on the micelle compositions. Most theories employ the pseudo phase separation approximation in combination with regular solution theory to model the intermicellar interactions responsible for the nonidealities in mixed micellar solutions.[207] The theoretical studies of mixed surfactant solutions available in the literature are mostly based on Rubingh's approach, because it includes a specific molecular interaction parameter β, defined as an interaction energy difference between same type of surfactant molecules.[95] Micellization in mixed surfactant solutions has been studied using a variety of experimental techniques (e.g., static and dynamic light scattering,[208] neutron scattering,[209–211] NMR[212]). However, mostly equimolar solutions were considered. Fang et al.[213] used the Fourier transform pulsed-field gradient (FT-PGSE) NMR technique to measure self-diffusion coefficients of TX100 and CTAB surfactants in mixed aqueous solutions at different molar ratios. Some important micellization parameters, such as CMCs of the mixtures, compositions of the mixed micelles and the interaction parameter β have been obtained. It was

demonstrated that at low surfactant concentrations deviations in the composition of the mixed micelles from the composition in the mixed solution occur. When the mole fraction of CTAB in the mixed surfactant solution, α_{CTAB} is higher than 0.5 micelles with a higher TX100 content were formed at first, whereas for α_{CTAB} lower than 0.5 micelles with a higher CTAB content were obtained. However, with an increase of the total surfactant concentration, the compositions of the mixed micelles approach the mole fraction of the mixed solution. The FT-PGSE NMR approach was also applied to study the micellization properties of Brij35/CTAB mixtures.[94] It has been shown that the mechanism of micelle formation is different from that of the TX100/CTAB system. At any composition of the mixed solution, Brij35 molecules have a tendency to form pure Brij35 micelles first and the CTAB molecules enter them with increasing total surfactant concentration. Nevertheless, in both systems with increasing surfactant concentration the composition of the mixed micelles approaches the composition of the solution. The composition of mixed Brij35/TX100 micelles at different ratios was estimated from NMR diffusometry self-diffusion coefficients and calculated with the regular solution approach by Koneva et al.[214] A change in the mixed micelle composition in dependence of the total surfactant concentration was observed. Higher amounts of Brij35 surfactants in the formed micelles were observed at concentrations near the CMC of the mixture, whereas at concentrations above 2 mM the composition of the micelles approaches the composition of the surfactant solution.

3.5 Partition equilibria of ionizable solutes in membranes and micelles

Many compounds of interest are ionizable and are charged in a specific pH range. Since the protonation equilibrium has a significant influence on the partition behavior, pH dependent partitioning of ionized solutes in various membrane/water systems is reported in the literature. Austin et al.[35] studied the partitioning of ionizable molecules in DMPC unilamellar vesicles as a function of pH using an ultrafiltration method. It could be demonstrated that octanol/water and membrane/water partition coefficients of unionized drugs have similar values, while those of ionized species have significant higher values in the lipid membrane system. The same behavior was observed for substituated phenols and methylated amines by Escher and Schwarzenbach[215] and Escher et al.[216] Therefore, it was concluded that liposomes constructed from zwitterionic phosphatidylcholines are appropriate model systems for mimicking the partitioning of both neutral and ionic molecules between biological membranes and water, while the

widely used octanol/water system is not suitable in case of charged solutes.[215] Avdeef et al.[217] determined liposome/water partition coefficients and lipophilicity profiles of ionizable drugs in DOPC unilamellar vesicles using a pH-metric technique. They found that the partition coefficients of the ionized drugs are lower than those of their neutral state due to electrostatic interactions with the zwitterionic membrane. The findings are consistent with those obtained with other methods such as EPR spectroscopy measurements, ultrafiltration and dialysis.[35,218–220]

The partition behavior of ionized solutes in micellar systems is less investigated in the literature. Mehling et al.[105] investigated the partition behavior of ionized solutes in Brij35 and CTAB micelles with MLC. Partition coefficients of dissociates acids in nonionic Brij35 micelles could not be determined due to the so-called non-binding behavior. As the dissociated solutes are more hydrophilic, they did not partition favorably in the nonionic micelles. Similar behavior was observed from cloud point experiments in aqueous micellar two phase systems, where very low partition coefficients of dissociated acids between TX114 surfactant rich phase and aqueous rich phase were obtained.[23] In CTAB micelles dissociated acids have higher partition coefficients than their protonated state due to enhanced electrostatic interactions with the cationic head group of CTAB.[105] Still, for ionized molecules, the determination of partition coefficients with retention models is challenging and in some cases not feasible due to very strong electrostatic interactions designated as overbinding.[105] The same behavior was obtained in mixed Brij35/CTAB micelles at higher CTAB content[42] and reported in previous study for the drug valsartan in CTAB micelles.[221] The prediction of partition equilibria of charged molecules in micellar systems is also challenging. There is a lack of theoretical methods able to predict the partition behavior of ionized solutes in micelles. However, predictions for charged solutes in lipid bilayer systems have been carried out already. As some of these concepts will be applied to micellar system in this thesis, they will be briefly reviewed in the next paragraphs.

COSMOmic for partition equilibria predictions of ionized solutes in membranes

It is known that the ordered structure of lipid bilayers results in an internal dipole potential, which cannot be measured directly. Several indirect approaches yield to hill-shaped potentials in the range from 227 mV to 280 mV in the membrane interior for DPPC bilayers and egg phosphatidylcholine bilayers.[222–224] Bittermann et al.[171] proposed an approach to improve the prediction quality of COSMOmic for charged molecules in membranes by implementing an

internal membrane dipole potential. They used an empirical model potential, a Gaussian-type error function with three adjustable parameters. The model was optimized to minimize the error in the prediction of a training set containing partition coefficients of ionic and neutral solutes in DMPC. The proposed refinement model improves the prediction quality of COSMOmic for ionic compounds, while it has no negative effect on the predictions of neutral molecules.[171] COSMOmic shows better prediction accuracy for ionizable molecules in membranes compared to models based on empirical correlations and is a promising tool for partition equilibria predictions of charged solutes.[225]

Thermodynamic cycle and position dependent pK_a

It is known that when both the neutral and the charged form of a solute bind to a lipid membrane, the pK_a of the solute in the membrane differs from that in the aqueous solution.[226,227] Shifts of ionization constants were observed in micellar systems as well.[221,228,229] In neutral lipids, the pK_a of acids increases and the pK_a of bases decreases due to the effect of the decreased dielectric constant in the interfacial zone.[36] In case the lipid vesicles or micelles are charged, an additional electrostatic shift occurs. If the surface is negatively charged it attracts protons into the interfacial zone, such that the interfacial pH will be lower than the bulk pH, therefore the apparent pK_a changes accordingly.[36] Ionization constants, interfacial electrostatic potential and protolytic equilibrium of acid-base indicators in various micellar pseudo phases are discussed in detail in ref 229. de Castro et al.[28] have shown that taking into account the protolytic equilibrium in the micellar phase is crucial when determining partition coefficients of ionizable drugs, as the partition behavior strongly depends on the ionization state of the molecule. Various theoretical methods including QSPR models, QM calculations and thermodynamic cycles were applied to predict pK_a values in both aqueous and organic solvent solutions.[230–233] Molecular models such as MD simulations have the advantage that they are able to estimate the change of the ionization constant when a molecule passes through different environments.[194] A common approach is to combine free energy profiles from MD simulations with a thermodynamic cycle to calculate the pK_a change as function of position in the lipid bilayer.[194] Chew et al.[234] used MD simulations to calculate free energy profiles of adamantanes in a POPC lipid bilayer and based on them to estimate the pK_a change within the membrane. It could be demonstrated that the protonated species (positively charged) partition favorably in the interfacial region, but have a large free energy barrier in the membrane center. Thus, adamantanes are distributed mostly in neutral form in the bilayer center. Based on atomistic MD simulations, MacCallum et al.[235] estimated the distribution

behavior and the position dependent pK_a of amino acid side chain analogs considering their neutral and charged form in a DOPC membrane. When charged molecules penetrate into the bilayer center, formation of water defects that hydrate the solutes was obtained. It was summarized that lysine, glutamine and aspartic acid partition favorably in their neutral state as the cost of forming a water defect in the membrane center is higher than the cost of protonating or deprotonating the solute. In contrast, for arginine it is energetically feasible to be located at the membrane center stabilized by a water defect. A favorable partitioning of arginine in charged form due to formation of water defect was reported in a DPPC bilayer as well.[236,237] It could be demonstrated that the combination of free energy profiles and thermodynamic cycles is a promising approach to predict position dependent pK_a values and therefore the ionization state of solutes within anisotropic systems such as lipid bilayers. The theory behind this concept is explained in more detail in chapter 4.5 and it is then applied to a micellar system in chapter 7.2.3.

3.6 Aim of this work

Most atomistic MD studies of micellar systems available in the literature investigate the self-assembly process with a focus on CMC prediction, as well as on the aggregates size and shape depending on the system conditions. However, only a few consider the partition equilibria of a third component (solute) between micelles and water on an atomistic scale. Most studies on partition equilibria predictions in surfactant-based systems apply the pseudo phase approach, which neglects the anisotropic structure of micelles and are restricted to partition equilibria of neutral molecules only. The aim of this work is to evaluate the combination between all-atom MD simulations and the COSMOmic model for solute partitioning prediction in all types (according to the surfactant head group charge) of single surfactant micelles (nonionic, anionic, cationic, zwitterionic) and in mixed nonionic/anionic and nonionic/cationic micelles. Factors, influencing the prediction quality of the approach, such as the size and the shape of the micelle structure obtained from MD simulations and used in COSMOmic calculations should be studied in detail.

The prediction of partition behavior in the nonionic surfactants Triton X-114 and Triton X-100 is of particular interest in this work due to their wide range of practical applications (e.g., in surfactant based extraction processes and membrane protein purifications). Since optimized parameters for these surfactants are not available in the known biomolecular force fields, in the first part of this thesis parameters for the Triton X series have to be optimized and validated enabling all-atom MD simulations. Regarding the micelle structures

obtained from self-assembly simulations, criteria to select suitable micelles for the COSMOmic calculations should be proposed. The partition equilibria prediction of neutral solutes in Triton X micelles should be evaluated, whereby the influence of the micelle structure used for the calculation with COSMOmic will be analyzed.

In the literature, the position dependent solute partitioning in micellar systems has been rarely investigated, whereby mostly one or two micelle types are considered. In this work, the favorable localization of selected solutes within various micelles should be estimated from free energy profiles. COSMOmic profiles and profiles calculated directly from MD simulations will be compared. Although the partition behavior of ionized solutes in micelles is of particular importance in pharmaceutical applications, most experimental and theoretical studies in the literature investigate only neutral molecules. There is a lack of theoretical methods to predict the partition behavior of ionized solutes in micelles. Therefore, in this work a special focus will be set on the description of charged solutes. Both MD simulations and COSMOmic should be evaluated for the prediction of ionized molecules. To make the validation of predictions possible, partition coefficients of dissociated solutes have to be obtained experimentally.

Despite the fact that for practical applications mostly mixtures of surfactants are used, the partition behavior in mixed micelles is less investigated in the literature compared to single surfactant systems. In this work, the micellization process and the composition of mixed nonionic/ionic micelles have to be studied with self-assembly MD simulations. The partition equilibria of neutral solutes in mixed micelles should be investigated, whereby COSMOmic predictions will be compared to COSMO-RS calculations using the pseudo phase approach. Furthermore, a thermodynamic cycle based on COSMOmic free energy profiles should be applied to estimate the position dependent pK_a value of a solute in nonionic/cationic micelles. Based on this, an approach to improve the accuracy of COSMOmic for partition coefficients prediction of ionizable solutes in mixed micelles should be proposed.

4 Materials and methods

4.1 Experimental determination of partition coefficients with micellar liquid chromatography

The surfactants Triton X-114 (TX114) and sodium dodecyl sulfate (SDS) were provided by Sigma-Aldrich and cetrimonium bromide (CTAB) was purchased from Serva Electrophoresis. The pH was adjusted with HCl from Th. Geyer. All chemicals were used without further purification.

Partition coefficients in pure CTAB micelles (chapter 6.1) and in mixed surfactant systems (equimolar solution of TX114/SDS and TX114/CTAB) (chapter 7.1.3) were determined using MLC. The chromatographic analysis was performed with an Agilent 1200 Series HPLC. The HPLC system included a quaternary pump, a tempered autosampler, a degasser, a diode array detector, and a column thermostat. A Nucleodur C18 Gravity column (Macherey Nagel, 4 x 125 mm, 5 μm, 100 Å) with the corresponding pre-column was used and kept at 298 K. The pH of the mobile phase was adjusted with NaOH and HCl to pH 2, pH 7 and pH 10.5. The mobile phase flow rate was set to 1 mL/min (for CTAB) and to 0.5 mL/min (for TX114/SDS and TX114/CTAB). The column was equilibrated with the surfactant solution until a constant pressure and constant UV signals have been reached. The column was loaded with surfactant until the adsorption equilibrium and a breakthrough curve was recorded. The retention of the solutes was determined for five different surfactant concentrations in the mobile phase (0.05-0.25 wt.% for CTAB and 0.1-0.5 wt.% for TX114/SDS and TX114/CTAB). The injection volume was 20 μL for all solutes. The partition coefficients were calculated according to the model of Armstrong and Nome.[101]

4.2 Analysis of micelles from MD simulations

4.2.1 Definition of micelles

In order to identify micelles from self-assembly simulations and to proof the stability of the preassembled micelles, the method proposed by Sammalkorpi et al.[181] was applied. In this procedure, usually three distances between surfactant monomers are defined, where two surfactant monomers are being part of the same micelle if (1) one of the three defined distances is shorter than r_{1cut}, or (2) two out of three distances are shorter than r_{2cut}, or (3) all three distances are shorter than r_{3cut}, where $r_{1cut} \leq r_{2cut} \leq r_{3cut}$. The distances and the cutoffs are chosen according to the specific systems after detailed visualization of the

configurations to verify a reliable definition of micelles. In the case of mixed micelles, two carbon atoms (one from each surfactant type) were grouped to calculate its distances. The distances and cutoffs used in this work are given in Table 4.1. In Figure 4.1 the structures of the surfactants are shown, where all atoms used for the definitions are highlighted.

Table 4.1: Distances between carbon atoms and cutoffs used for the definition of micelles. The atoms of the nonionic surfactant are given first (e.g., in the last row, column two C5 of TX114 and C19 of CTAB). In some cases the center of mass (COM) of a surfactant is used instead of an atom.

Surfactant system	Distances			Cutoffs [nm]		
				r_{1cut}	r_{2cut}	r_{3cut}
Brij35	C1	C6	C12	0.68	0.68	1.20
TX114;TX100	C1D	C1A	C5	0.58	0.78	1.00
HePC	C6	C15	C21	0.66	0.88	1.10
SDS	SDS COM	C5	C12	0.61	0.66	0.76
CTAB	CTAB COM	C8	C19	0.70	0.74	0.89
Brij35/CTAB	C12 + C19	C6 + C18	C1 + C16	0.68	0.68	1.20
TX114/SDS	C5 + C12	C4 + C11	C2 + C10	0.58	0.78	0.88
TX114/CTAB	C5 + C19	C4 + C14	C1D + C6	0.58	0.78	1.00

4.2.2 Shape analysis

The shape of micelles has been analyzed by examining the eccentricity ε,[238,239] which is defined as

$$\varepsilon = 1 - \frac{I_{\text{min}}}{I_{\text{avg}}} \tag{4.1}$$

where I_{min} is the moment of inertia along the principal axes with the smallest magnitude, whereas I_{avg} is the average of all three moments of inertia. ε can have values from 0 to 1, inclusive. For simple, closed aggregates, $\varepsilon = 0$ usually corresponds to a highly symmetrical shape resembling a sphere while larger values of ε correspond to progressively more elliptical shapes.[238,239]

The radius of gyration of the micelles is calculated as the root mean square distance between the center of mass of the micelle and its ends. If the micelle is a sphere, the radius of this sphere R_{s} can be calculated from the relation to the radius of gyration

$$R_{\text{s}} = \sqrt{\frac{5}{3}} R_{\text{g}} \tag{4.2}$$

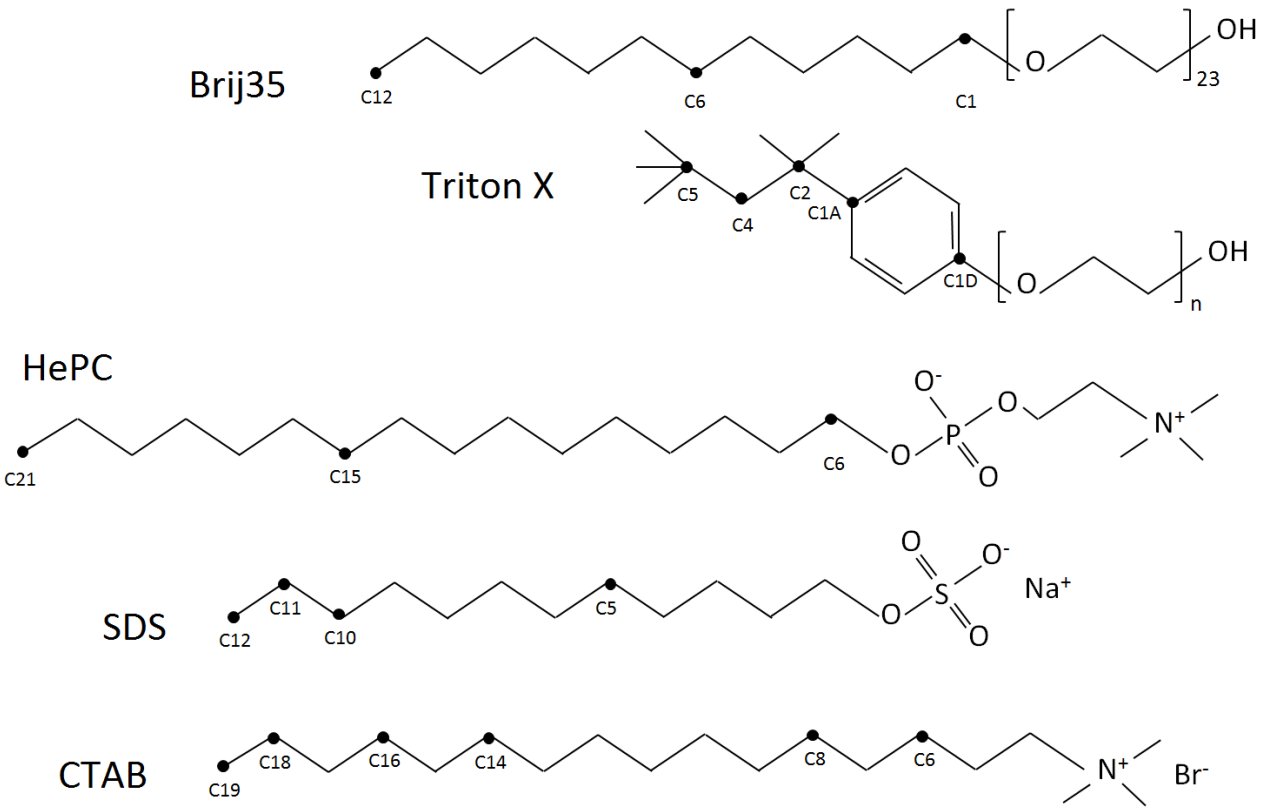

Figure 4.1: Structures of the surfactants: Brij35 (nonionic), Triton X (nonionic), HePC (zwitterionic), CTAB (cationic), SDS (anionic). The carbon atoms used for the micelle definition are highlighted. In case of Triton X, the hydrophilic chain lenght n would define the specific Triton X (e.g., TX100 or TX114).

4.3 Averaged atomic distributions for COSMOmic

To predict the micelle/water partition behavior of solute molecules with COSMOmic, a micelle structure is required. Since the model COSMOmic needs the spatial composition of the system, the use of proper micelle structures is crucial for the predictive quality. In the study of Storm et al.[185] single micelle structures were used as input for the COSMOmic calculations, whereas thousands of single micelles have been selected in order to achieve statistical reliable results. More physically reasonable approach is to use averaged atomic distributions as input structures for the COSMOmic calculation.[240] In this work, averaged atomic distributions of the most probable micelle sizes, which have been obtained during the MD simulations, have been calculated. This procedure was previously suggested for bilayers and it was demonstrated that using averaged atomic distributions from MD simulations as input for the model COSMOmic is beneficial.[240] In order to achieve statistically reliable results, the atomic distribution were averaged over at least 200 micelles.

It has been demontrated that gaps in the water layer around some micelles can occur, resulting from another micelle in close proximity.[170] When using such micelle structures as input for COSMOmic, a decline in the prediction quality was observed. In order to overcome the effect of outliers, resulting from a deficient water layer around the micelle, an approach to fill up the gaps in the water shell was applied in this work. If atom i of a surfactant molecule, which is not part of the micelle, was present in the atomic distribution, it was subsequently replaced by an equivalent amount of water molecules. This amount was calculated by

$$N_{\text{sol,fill}} = \frac{n_{\text{sol}}}{n_{\text{surf}}} \frac{m_{\text{surf},i}}{M_{\text{surf}}} \tag{4.3}$$

where n is the molecular bulk density, M the molecular weight, and m the mass of an atom. The subscripts "surf" and "sol" denote the surfactant and solvent molecule, respectively. It is important to note that only the values of the water molecules were altered by this method, no atoms were placed into space. In order to validate this approach, the radial density of water around a micelle was analyzed (TX114 micelle with $N_{\text{agg}} = 33$). The micelle structure was obtained from a self-assembly simulation. The same micelle was also separately simulated in a system containing otherwise only water. Therefore, a snapshot of the self-assembly simulation, containing a $N_{\text{agg}} = 33$ micelle were taken and all other Triton X-114 molecules were removed. After a short equilibration simulation, the system was simulated for 40 ns. In this case, no gaps in the water layer can occur, as there is only one micelle in the system. Therefore, the radial distribution of water from this system corresponds to a complete water layer. As COSMOmic assumes the second phase for the partition coefficient calculations to be the last layer, the densities far away from the micelle should represent bulk water for micelle/water partition coefficient predictions. Hence, the densities around the different micelles should level off at the same water density. However, by comparing the densities for the micelle from the self-assembly simulation with the density of the micelle surrounded by water only, it can be seen that they converge to different values (see Figure 4.2). This is due to other micelles and surfactant molecules in the self-assembly simulation. When the gaps are filled with the described approach, the radial density overlap with the density from the simulation of a single micelle. Therefore, it can be concluded, that the applied approach successfully generates structures with complete water shells around the micelle. In this work, all atomic distributions are calculated with this approach.

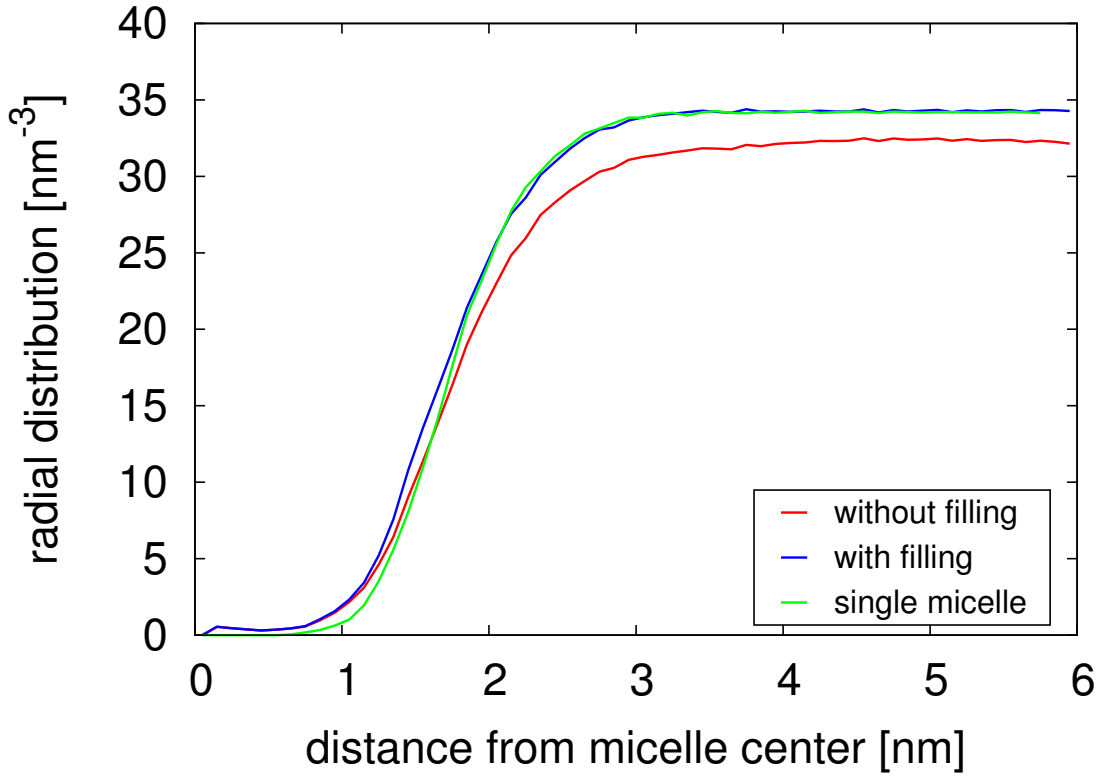

Figure 4.2: Radial density of the oxygen atom of water around a TX114 micelle: calculated from a self-assembly simulation without filling up the gaps in the water shell (red); calculated from the same self-assembly simulation, whereas the gaps in the water shell were filled up (blue); calculated from a simulation of a single micelle in a water box (green).

4.4 COSMOmic for free energy profiles in micelles

In the case of micelles the volume of each individual shell depends on the distance between the bulk water layer and the center of the micelle. Therefore, when using a constant layer thickness, the volume of the layers increases with this distance. This has an influence on the shape of the free energy profile. As the free energy is a measure for residence probability, decreasing free energy (increasing probability) is calculated when the volume increases. Therefore, the free energy in the water phase is decreasing instead of reaching a constant value. To overcome this effect, the free energy can be calculated with corrected volumes of the layers (see equation 2.60). In Figure 4.3, a free energy profile with corrected volumes is compared to an unmodified COSMOtherm result. It is important to understand the different meaning of the profiles. The unmodified profiles give the free energy of the bin. That is taking into account different volumes the free energy is different in the same environment (e.g., water) (in a larger volume the probability is higher to find a solute). In contrast, the volume corrected profile represents the free energy due to the environment (interactions with surrounding molecules). All profiles, shown in this work, are calculated

with corrected volumes.

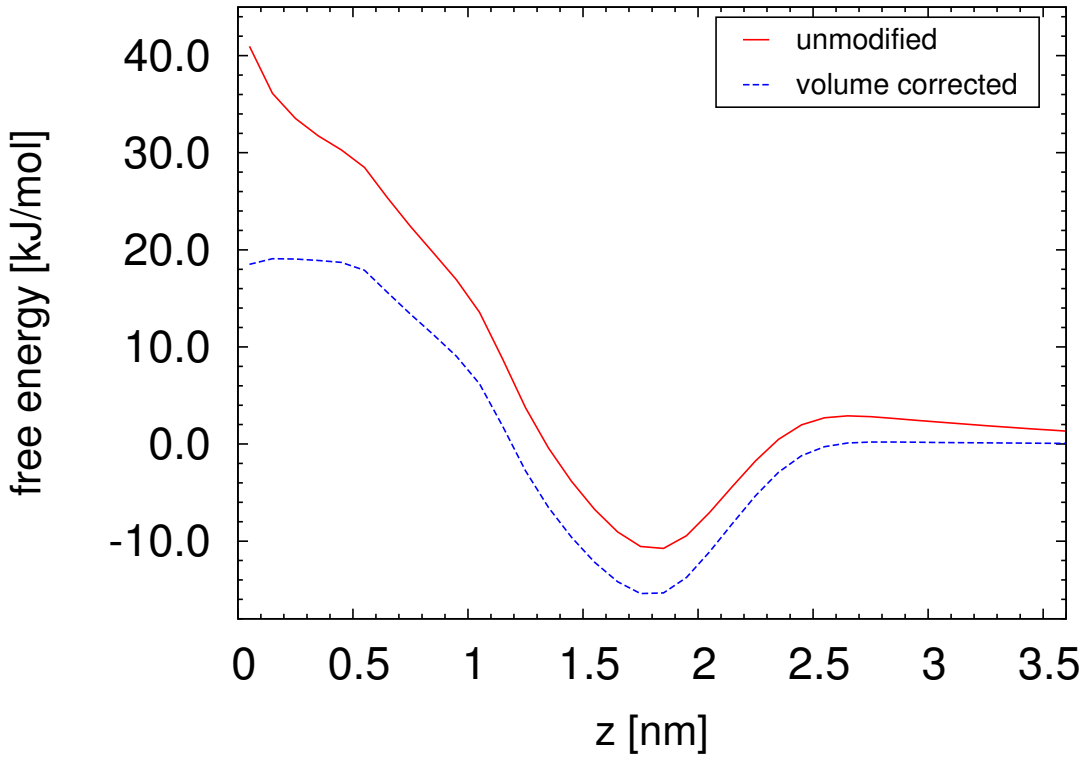

Figure 4.3: Free energy profiles of 4-hydroxybenzoic acid in a CTAB micelle. The profile shown in blue is calculated with weighted volumes of the shells, the profile shown in red is an unmodified COSMOtherm result.

4.5 Thermodynamic cycle and position dependent pK_a

It is known that when both the charged and the uncharged forms of a drug bind to a membrane with different binding constants, the pK_a for the drug in the membrane should differ from that in the aqueous phase.[226] This change of the pK_a compared to the bulk aqueous phase is obtained for micelles as well and has to be taken into account.[28,241] The position dependent pK_a as a function of the distance from the micelle center can be estimated based on the free energy profiles of both the charged and neutral form of the molecule. The thermodynamic cycle, which allows this calculation is given in Figure 4.4. ΔG_{AH} and ΔG_{A^-} are the free energy profiles of the neutral and dissociated molecule, in this work they were calculated with COSMOmic. The free energy of dissociation in water ΔG_{AHw} at a given pH can be calculated with

$$\Delta G_{AHw} = 2.303RT(\mathrm{p}K_a - \mathrm{pH}) \tag{4.4}$$

The described procedure below is applied for the partition coefficients calculation of ionized isovanillin in section 7.2.3. From these three terms the fourth

term, the free energy of dissociation in the micelle ΔG_{AHm} can be calculated (see Figure 4.4) and therefore the pK_a in the micelle. In this calculation, the protolytic equilibrium of the hydronium ion between the micelle and water is neglected.[242]

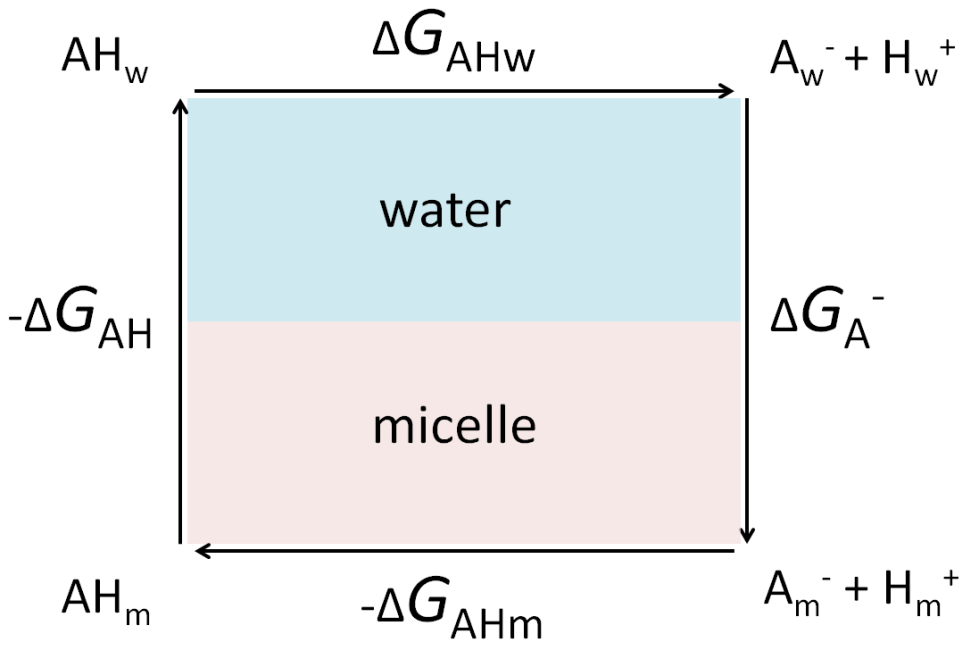

Figure 4.4: Thermodynamic cycle to calculate the pK_a as a function of distance from the micelle center. This pK_a is related to the free energy of dissociation in the micelle ΔG_{AHm}. ΔG_{AH} and $\Delta G_{\mathrm{A^-}}$ are the free energies for the transfer from the water to the micelle phase for the neutral and dissociated solute. ΔG_{AHw} is the free energy of dissociation in the water phase (see equation 4.4).

To calculate partition coefficients of ionizable solutes both the protonated (AH) and dissociated (A^-) form of the molecules have to be considered.[28] Each form is distributed between the micelle and the aqueous phases according to their specific partition coefficients (K_{AH} and $K_{\mathrm{A^-}}$). The total partition coefficient that accounts for the partitioning of both forms K_{T} can be expressed as the weighted average of K_{AH} and $K_{\mathrm{A^-}}$ according to[226]

$$K_{\mathrm{T}} = x_{\mathrm{AH}} K_{\mathrm{AH}} + x_{\mathrm{A^-}} K_{\mathrm{A^-}} \tag{4.5}$$

where x_{AH} and $x_{\mathrm{A^-}}$ are the mole fractions of the protonated and dissociated form in the micelles, respectively. x_{AH} and $x_{\mathrm{A^-}}$ are calculated as average over the micelle layers, whereas the mole fractions at different positions (layers) in the micelle are calculated at the desired pH (e.g., in chapter 7.2.3 it was pH = 10.5 as used in the corresponding experiments[42]) from the position dependent pK_a with the Henderson-Hasselbalch equation (see equation 2.28, chapter 2.2.3) under the condition $x_{\mathrm{AH}} + x_{\mathrm{A^-}} = 1$. K_{AH} is the partition coefficient of the

neutral form at the given Brij35/CTAB micelle composition calculated with COSMOmic.

It is assumed that the partitioning of dissociated solutes in the nonionic surfactant can be neglected.[42,243] Hence, K_{A^-} is the partition coefficient of the dissociated form in pure CTAB micelle calculated with COSMOmic.

4.6 Error calculation

In this work, deviations between experimental and predicted partition coefficients were estimated with the root mean square error (RMSE), defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log K_i^{\mathrm{calc}} - \log K_i^{\mathrm{exp}})^2} \tag{4.6}$$

where n is the number of solutes for which calculations were carried out.

5 Modeling of Triton X-114 and Triton X-100 micelles

In this chapter the aggregation process of the nonionic surfactants Triton X-114 (TX114) and Triton X-100 (TX100) is investigated with self-assembly MD simulations. As the fundamental models behind every MD simulation are the force fields, the quality of the force field parameters is of major importance for the reasonable representation of physical properties with MD simulations. Therefore, at first force field parameters for the surfactants of the Triton X series are optimized and validated. Furthermore, the self-aggregation process of TX114 and TX100 is simulated at different concentrations and temperatures. In section 5.2.3, micelles obtained from the self-assembly simulations are used for prediction of partition equilibria with COSMOmic, whereas the influence of micelle size and shape on the prediction quality is investigated. The results of the following sections were in parts published as Yordanova, D.; Smirnova, I.; Jakobtorweihen, S. **Molecular Modeling of Triton X Micelles: Force Field Parameters, Self-Assembly and Partition Equilibria** in Journal of Chemical Theory and Computation.[173]

5.1 Optimization of force field parameters

5.1.1 Computational details

The surfactant TX114 is not directly included in the CHARMM36 force field nor in the CHARMM General Force Field (CGenFF). Mayne et al.[143] developed the Force Field Toolkit (ffTK), which is distributed as a VMD plugin and facilitates an organized workflow with a graphical user interface design for rapid parametrization of CHARMM compatible parameters. The ffTK (Version 1.0)[143] within VMD 1.9.2 alpha release, was used for the optimization of the missing parameters. The general parametrization procedure is summarized in chapter 2.3.1 and fully described elsewhere.[112,143] The optimization is based on performing molecular mechanics (MM) calculations which should reproduce the quantum mechanics (QM) target data.[112] All QM calculations were performed using Gaussian03.[244] Initial guesses were obtained with the ParamChem tool.[144,145] In addition, a "penalty score" is assigned by ParamChem to each parameter which judges the quality of the parameters. According to these penalties eight atom charges, one bond, four angles and nine dihedrals were parametrized in this work, the penalties of these parameters were between 1 and

16. All other parameters were used from CGenFF version 2b7 as their penalties were zero. For further details about the ParamChem algorithm and the penalty score estimation, please refer to the corresponding publications.[144,145] CGenFF LJ parameters were assigned to existing atom types. The molecular geometry was optimized at the MP2/6-31G(d) level of theory.

Charge optimization was based on compound-water interactions, which were built for all hydrogen bond donors. The interaction distance was optimized at the HF/6-31(d) level of theory, keeping all other degrees of freedom fixed. The interaction energy was scaled by a factor 1.16 and the QM hydrogen bond length was offset by -0.2 Å to be relevant for the bulk phase.[112] The charge optimization was run iteratively in simulated annealing mode.[143]

The QM target data for the optimization of the bonded parameters was obtained from the Hessian computed at the MP2/6-31G(d) level of theory and scaled by 0.89.[112] The optimization was iteratively performed in downhill mode[143] weighting the geometry and energy terms 1:1.

A potential energy scan (PES) for each selected dihedral angle was calculated in 15° increments, starting from the optimized geometry. Two Gaussian input files were generated for each scan - one in the positive direction and one in the negative direction. Each structure was optimized at the MP2/6-31G(d) level of theory. The dihedral optimization was based on reproducing the QM PES. The first round of optimization was performed in simulated annealing mode. Several refinement optimizations were run, again in simulated annealing mode. A final refinement was performed in downhill mode.[143]

5.1.2 Parameter optimization overview

CGenFF parameters for TX114 molecule were obtained using the ffTK.[143] In line with the CHARMM force field philosophy,[112] two smaller molecules representing the hydrophobic part of TX114 were used for the parameter calculations. The parameters were first optimized for the smaller molecule in order to get better initial guesses for subsequent iterations. Schematic illustration of the TX114 structure and the two smaller molecules are presented in Figure 5.1. Note that when a covalent link is formed to create a larger molecule from model compounds, the charge on the deleted H is added to the adjacent oxygen to maintain an integer charge. The parametrization was carried out by applying the following sequence:

1. Optimization in four steps, where first charges, second bonds, then angles and in the end dihedrals of model compound (1) (see Figure 5.1b), are obtained.

2. Second iteration of Step 1.

3. Reoptimization of C2, C4 and C5 charges, bonds, angles and dihedrals in model compound (2), for atom names and molecule structures see Figures 5.1c and 5.2.

4. Second iteration of Step 3 until reaching convergence to reproducible values.

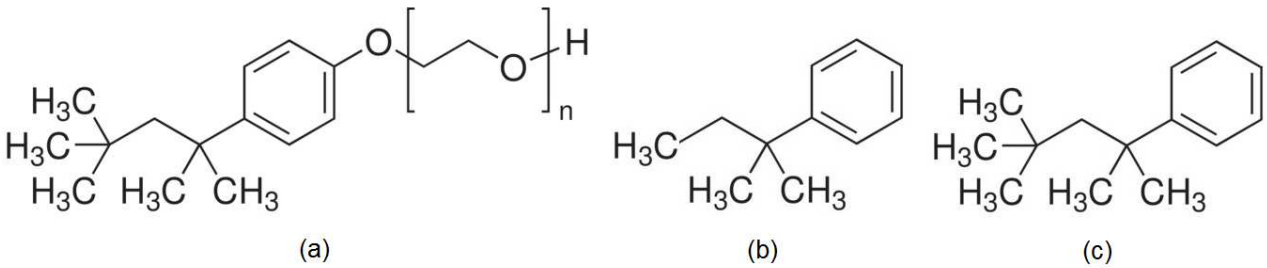

Figure 5.1: (a) Structural formula of TX114 (n=7-8) and TX100 (n=9-10); in the performed MD simulations TX114 and TX100 were modeled with n=8 and n=10, respectively. (b) Structural formula of model compound (1): 1,1-dimethylpropyl-benzene; (c) structural formula of model compound (2): 1,1,3,3-tetramethyl-benzene.

Partial charges were optimized for six carbon atoms: C1A, C2, C3A, C3B, C4 and C5 (see Figure 5.2). According to the CHARMM force field,[112] the charges of the nonpolar hydrogens were assigned a fixed charge of +0.09. Aromatic C-H groups were assigned the standard CHARMM charges of -0.115 and +0.115 on carbon and hydrogen atoms, respectively. The charges of the three carbon atoms that are not optimized are taken from the CGenFF (they show a penalty of zero in the ParamChem tool). QM calculations for interactions between the model compounds and water molecules have been computed following the standard CGenFF procedure. The interaction properties of the final optimization are listed in Table 5.2. The optimized partial charges are given in Table 5.1.

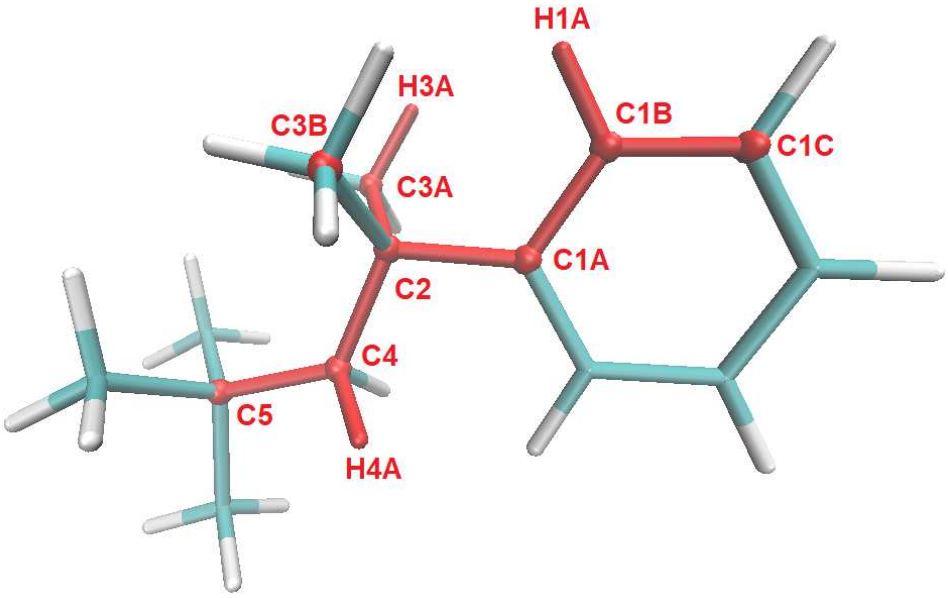

Figure 5.2: 3D structure of 1,1,3,3-tetramethyl-benzene (model compound (2)); the atoms for which the atomic charges were optimized and the atoms comprising the eight dihedral angles which were optimized in this molecule are shown in red.

Table 5.1: Optimized atomic charges.

Atom	Atom type	Optimized charge
C5	CG301	-0.096
C4	CG321	-0.085
C2	CG301	0.162
C3A	CG331	-0.265
C3B	CG331	-0.265
C1A	CG2R61	-0.172

Ideally, the compound-water energies should be within 0.2 kcal/mol from the QM interaction energies.[112] Some of the final MM energies differ by more than 0.2 kcal/mol from the target data, but the agreement is satisfactory considering that the matching of all 22 interaction energies is complicated for a large molecule like model compound (2). Deviations in this range are also reported for optimized CGenFF parameters for phosphonosulfonimidate.[245] The interaction distances are overestimated by 0.1-0.3 Å(see Table 5.2). As shown in Table 5.3, the dipole moment is overestimated by 47% compared to the MP2 value, which is in the desirable range to reproduce bulk phase properties.[112]

Table 5.2: Optimized 1,1,3,3-tetramethyl-benzene (model compound (2))-water interaction energies and geometries.

	Interaction energy (kcal/mol)	Interaction distance (Å)
Interaction geometry	ΔE(QM − MM)	Δr(QM − MM)
H1A...OHH	-0.385	-0.050
H1B...OHH	0.503	0.250
H1C...OHH	0.585	0.300
H1D...OHH	0.618	0.300
H1E...OHH	0.545	0.300
H3A...OHH	0.400	0.150
H3B...OHH	0.128	0.300
H3C...OHH	0.165	0.150
H3D...OHH	0.165	0.150
H3E...OHH	0.400	0.150
H3F...OHH	0.128	0.300
H4A...OHH	-0.052	0.150
H4B...OHH	-0.052	0.150
H6A...OHH	0.094	0.100
H6B...OHH	0.094	0.100
H6C...OHH	0.527	0.200
H6D...OHH	0.196	0.150
H6E...OHH	0.239	0.250
H6F...OHH	0.524	0.200
H6G...OHH	0.239	0.250
H6H...OHH	0.196	0.150
H6I...OHH	0.476	0.200
AD[a]	0.249	0.182
AAD[a]	0.317	0.204
RMSD[a]	0.358	0.209

[a]AD, average deviation; AAD, average absolute deviation; RMSD, root-mean-square deviation.

Table 5.4 represents the optimized bonded parameters along with the QM results. QM bond length is reproduced within 0.3 Å, as desirable.[112] MM optimized angles are all within 3°, except for C2-C4-C5 angle which shows a deviation of 10°. However, the result is considered as reasonable, because it is in the range of the CGenFF value (113.50°), for which the ParamChem tool shows a penalty as low as 1.8. The optimized bonds and angles parameters are given in Table 5.5.

Table 5.3: Gas phase dipole moment of model compound (2) calculated at the MP2/6-31G(d) and MM level of theory.

μ	QM	MM
X	-0.305	-0.381
Y	-0.055	-0.249
Z	0.000	0.000
Total	0.310	0.456

Table 5.4: QM and MM optimized equilibrium geometry of model compound (2).

Coordinate	QM	MM	Difference
Bond lenghts (Å)			
C1A-C2	1.5426	1.5420	-0.0006
Angles (°)			
C1B-C1A-C2	121.58	121.80	0.22
C1A-C2-C4	108.80	108.40	-0.4
C1A-C2-C3A	107.56	108.70	1.2
C2-C4-C5	126.24	116.24	-10.0

Table 5.5: Optimized bonds and angles.

Bonds			
Coordinate	Atom type	k_b [a]	b_0 [b]
C1A-C2	CG2R61-CG301	288.00	1.5420
Angles			
Coordinate	Atom types	k_θ [c]	θ_0 [d]
C1B-C1A-C2	CG2R61-CG301-CG301	30.96	121.80
C1A-C2-C4	CG2R61-CG301-CG321	119.37	108.40
C1A-C2-C3A	CG2R61-CG301-CG331	33.55	108.70
C2-C4-C5	CG301-CG321-CG301	47.97	116.24

[a] k_b, bond force constant
[b] b_0, equilibrium bond length
[c] k_θ, angle force constant
[d] θ_0, equilibrium angle

Table 5.6: Optimized dihedral angles.

Coordinate	Atom types	k_Φ [a]	n [b]	δ [c]
C1C-C1B-C1A-C2	CG2R61-CG2R61-CG2R61-CG301	3.000	2	180.00
C2-C1A-C1B-H1A	CG301-CG2R61-CG2R61-HGR61	2.172	2	180.00
C1B-C1A-C2-C4	CG2R61-CG2R61-CG301-CG321	1.770	2	180.00
C1B-C1A-C2-C3A	CG2R61-CG2R61-CG301-CG331	2.333	2	180.00
C1A-C2-C4-C5	CG2R61-CG301-CG321-CG301	0.1030	3	0.00
C1A-C2-C4-H4A	CG2R61-CG301-CG321-HGA2	0.3230	3	0.00
C3A-C2-C4-C5	CG331-CG301-CG321-CG301	0.1730	3	0.00
C1A-C2-C3A-H3A	CG2R61-CG301-CG331-HGA3	0.1880	3	0.00
O2-CB1-CB2-O3	OG301-CG321-CG321-CG311	0.8720	1	180.00
O3-CC1-CC2-O4	OG301-CG321-CG321-CG311	0.9730	2	0.00

[a] k_Φ, dihedral force constant
[b] n, multiplicity
[c] δ, phase shift

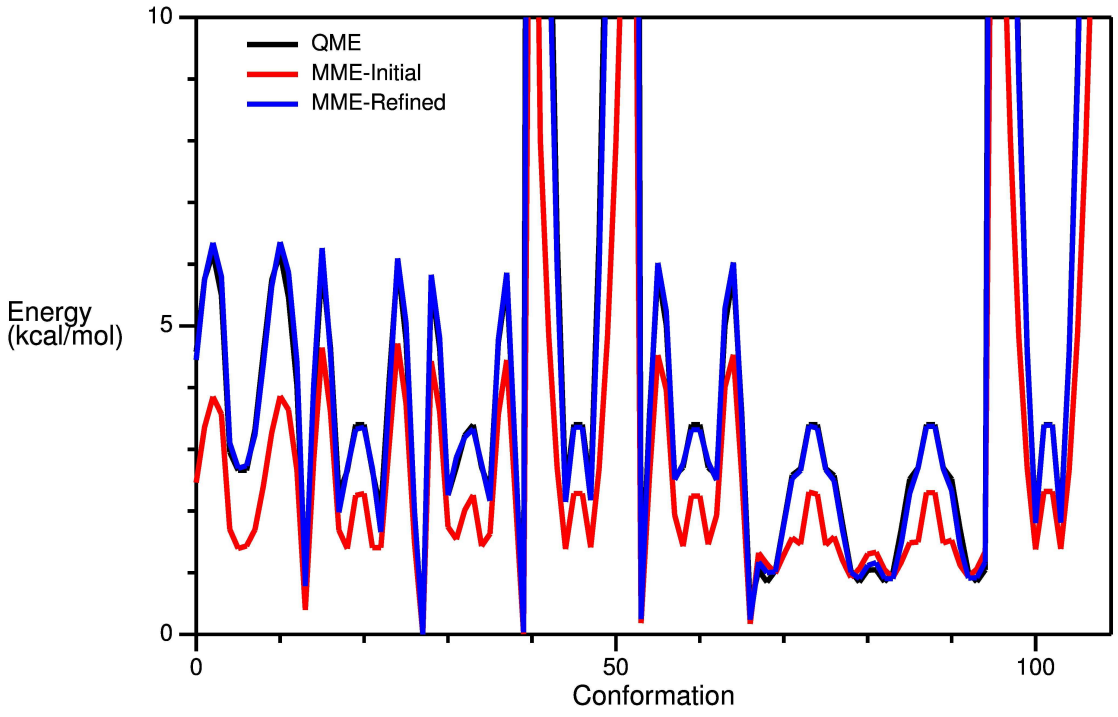

Figure 5.3: The potential energy scans for the eight optimized dihedral angles of model compound (2), shown in one plot. The abscise depicts the number of the scan point. The ordinate is the relative energy in kcal/mol: QM PES (black); using the initial guess parameters (red); using the optimized parameters (blue). The results were visualized with ffTK.[143]

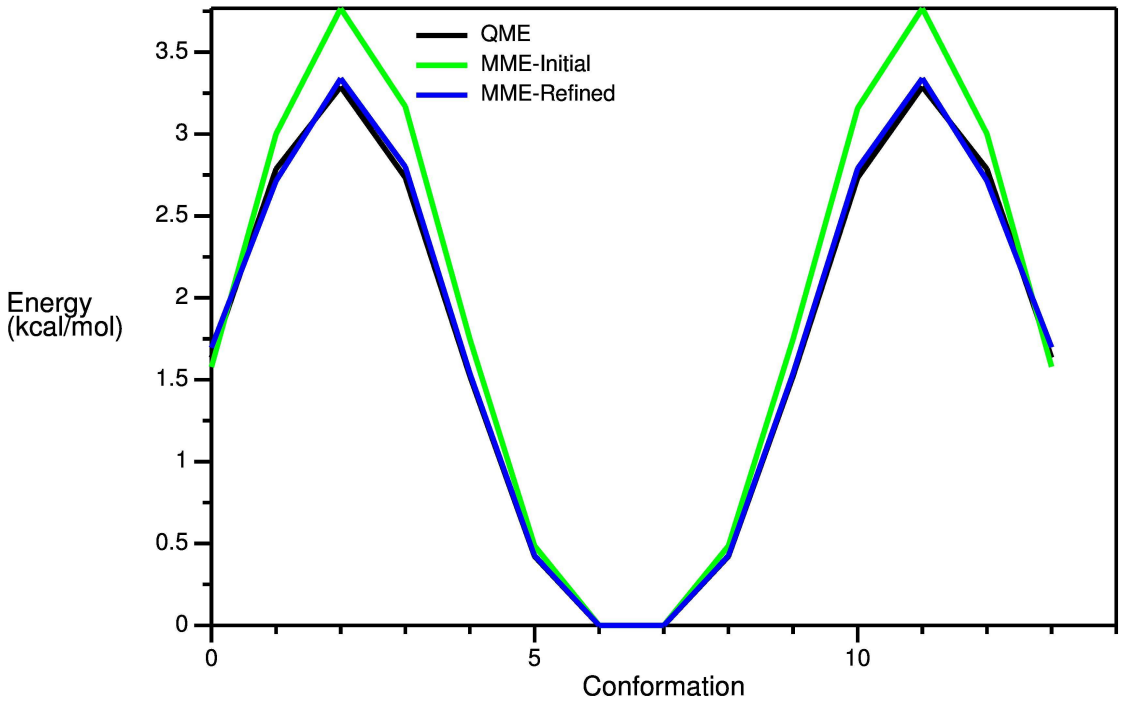

Figure 5.4: Potential energy scans of one optimized dihedral from the hydrophilic poly(ethylene oxide) chain. The abscise depicts the number of the scan point. The ordinate is the relative energy in kcal/mol: QM PES (black); using the initial guess parameters (green); using the optimized parameters (blue).

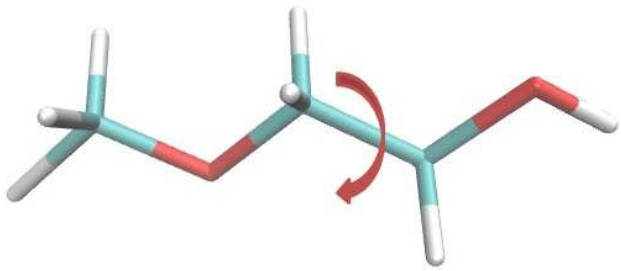

Figure 5.5: Structure of ethylene glycol monomethyl ether used for optimization of one dihedral from the hydrophilic head region of TX114.

Figure 5.3 shows a comparison between the target QM, initial MM and refined MM PESs of eight optimized torsion angles (see also Figure 5.2). All optimized dihedral angles are given in Table 5.6. The refined MM parameters are in excellent agreement with the QM PESs. The RMSE between the QM and the refined MM PESs for all eight dihedral angles is 0.125. One dihedral from the hydrophilic poly(ethylene oxide) chain (hydrophilic head) of TX114 was separately optimized, ethylene glycol monomethyl ether was used for this optimization (see Figure 5.5). The refined MM parameters yield a PES that reproduces the QM target PES with excellent agreement (RMSE = 0.044).

A comparison between the MP2/6-31G(d), initial MM and refined MM PESs is shown in Figure 5.4.

Validation

To validate the determined force field parameters usually some physical and thermodynamic quantities (density, enthalpy of vaporization, energy of solvation) are calculated and compared with experimental measurements.[112,143] For that, a pure TX114 bulk liquid phase has to be simulated. A rectangle box containing 450 TX114 molecules was constructed by placing a copy of the molecule on each grid point of a rectangle 2.625x10.5x15 lattice. After a energy minimization with the steepest descent method, a short simulation in the NPT ensemble was conducted for 400 ps, using the Nosé-Hoover thermostat[246] with the coupling time constant of 1 ps at 298.15 K and the Berendsen barostat[247] (isotropic, coupling constant $\tau_p = 5$ ps). As the volume was far from equilibrium and for a rapid equilibration, the pressure was set to 500 bar in this short equilibration step. Afterwards, the equilibrated box has been simulated in the NPT ensemble for 12 ns, using the Nosé-Hoover thermostat[246] at 298.15 K and the Parrinello-Rahman barostat[248] at 1 bar. From this simulation the density of TX114 was determined. This (1.055 g/cm^3) is in excellent agreement with the experimental value of 1.052 g/cm^3.[23] Unfortunately, experimental data for other pure TX114 properties, which can be used for parameter validation, are not available.

5.2 Self-assembly in aqueous solution and partition equilibria

5.2.1 Simulation details

The MD simulations were performed using the GROMACS package version 4.6.5.[249] To obtain initial configuration for the MD simulations, a specific amount of TX114 monomers was placed randomly into a cubic water box to match a specific concentration, whereby overlapping water molecules were removed. The energies of the initial configurations were minimized with the steepest descent method. After the energy minimization, a simulation in the NVT ensemble was conducted for 600 ps to further relax the system, using the Nosé-Hoover thermostat[246] with the coupling time constant of 1 ps and $T = 298$ K. After these initialization steps, all simulations were performed in the NPT ensemble at $p =$ 1 bar and $T = 283$ K or $T = 313$ using the Nosé-Hoover thermostat[246] (coupling constant $\tau_t = 1$ ps) and the Parrinello-Rahman barostat[248] (coupling constant $\tau_p = 2$ ps). For the NPT simulations a time step of 0.002 ps has been used. The Lennard-Jones interactions were cut off at 1.2 nm and switched from 0.8 nm on, where both the potential and the force were switched. For the Coulomb interactions, a real space cutoff of 1.0 nm was applied. For long-ranged electrostatic interactions the Particle Mesh Ewald method[133,250] was used. All bonds to hydrogens were constrained with the LINCS algorithm.[251]

The self-assembly of TX114 was studied at two different concentrations (0.1 mol/L and 0.22 mol/L) and two different temperatures (283 K and 313 K). Additionally, a simulation of TX100 self-assembly at 0.22 mol/L and 283 K was performed. A summary of all MD simulations is given in Table 5.7.

Table 5.7: System configurations for MD simulations of TX114 and TX100 in aqueous solution.

	Number of molecules					
Surfactant	Surfactant	Water	Total number of atoms	$C_{\text{surfactant}}$ [mol/L]	Temperature [K]	Simulation time [ns]
TX114	216	53352	180144	0.22	283	100
TX114	216	53371	180201	0.22	283	100
TX114	216	53360	180168	0.22	313	100
TX114	216	53360	180168	0.22	283	200
TX114	216	120515	381633	0.1	283	200
TX100	216	53407	183333	0.22	283	200

COSMO-RS and COSMOmic

The COSMOmic calculations require at least one conformer for each molecule type in the system. As the conformation of amphiphilic molecules is dependent on the environment, surfactant monomers were taken from MD simulations to obtain conformers present in micelles. Representative surfactant monomers were selected with respect to the solvent accessible surface.[240] The surfactant conformers, as well as the solute conformers were calculated with Turbomole 5.10[252] on the BP-TZVP[253–255] level and the RI (resolution of the identity) approximation.[256] A conformer analysis of the solute molecules has been performed with HyperChem 8.0.[257] In this chapter, the COSMOtherm (version C3.0 Release 13.01) implementation of COSMO-RS and COSMOmic with the BP_TZVP_C30_1301 parametrization was used.

5.2.2 Self-assembly: Influence of concentration and temperature

MD simulations of TX114 and TX100 in aqueous solution have been performed. The simulated concentrations (0.1 mol/L and 0.22 mol/L) are both orders of magnitude higher than the CMCs of TX114 and TX100 (0.17 mM[258] and 0.24 mM[259], respectively). Lower concentrations were not simulated due to limitations of the system size. The simulated systems are already very large (see Table 5.7), consequently computationally very demanding. However, at the considered concentrations the TX114/water and TX100/water systems are both in the one phase region (isotropic liquid).[260,261]

In Figure 5.6a the progress of the maximum aggregation number, N_{agg} over the simulation time at two different concentrations of TX114 is shown. The aggregation numbers of the observed micelles vary between 30 and 216 over the MD trajectories, which corresponds to the growing micelles sizes during the equilibration process and to the polydispersity of nonionic surfactants solutions.[175,262,263] In the system at concentration 0.22 mol/L, small aggregates consisting of 30-40 surfactant molecules were formed after 10 ns. In the first 100 ns of the simulation micelles in close proximity were observed (see Figure 5.7). The assembly of small aggregates into larger micelles ($\sim$100 surfactant monomers) starts at 40 ns. The maximum N_{agg} after 100 ns is around 120. At 120 ns, all surfactants begin to aggregate into one micelle (216 surfactant monomers). At the lower concentration of 0.1 mol/L small micelles with $N_{\mathrm{agg}} \approx 30$ were observed after 25 ns. The maximum N_{agg} after 50 ns is around 60 and after 100 ns smaller aggregates start to aggregate to a larger micelle with an aggregation number of 183. The maximum N_{agg} over the 200 ns MD trajectory remains 183 (lower than 216 at the higher concentration of 0.22 mol/L).

Due to the fact, that at the lower concentration two times more water is present, the micellization process is slower and in general smaller aggregates are observed. Faster micelle formation with increasing concentration is reported in other MD studies, as well.[181,185] However, the equilibrium condition in these systems may not have been reached. Since at the lower concentration the aggregation process is slower, it can be assumed that longer simulation times are needed for this system to reach complete equilibration. Therefore, it is expected that with enough simulation time both systems would converge to a single large micelle containing all 216 monomers. Wolszczak and Miller[264] reported experimentally determined aggregation numbers of 68 and 156, dependent on the used fluorometric techniques. The measurements were carried out at TX114 concentration of 0.01 mol/L. It is known that aggregation numbers determined at surfactant concentration higher than the CMC may differ strongly from the aggregation numbers around the CMC.[258] The concentrations in the experimental determinations and in the simulations are both orders of magnitude above the CMC, but also differ by an order of magnitude from each other. However, aggregation numbers of nonionic surfactants increase with increasing concentration.[59,60] Therefore, the aggregation numbers in the simulations should be higher than those in the experiments by Wolszczak and Miller[264] which corresponds to the findings.

In order to investigate the influence of temperature on the aggregation process, an additional 100 ns simulation of 0.22 mol/L TX114 solution with 216 surfactant molecules has been performed at a temperature above the CPT of TX114, which is ∼300 K at this concentration.[243] In Figure 5.6b, the growth of the micelles in terms of the largest aggregation number is presented. Corresponding to experimental results, at temperatures above the CPT the micellization process is faster and the micelle size increases significantly.[265] The simulations are qualitatively in agreement with these findings, as self-assembly is faster at 313 K and the largest micelles containing all 216 surfactant molecules are already stable after 50 ns. As shown in Figure 5.6a at 283 K the micelle size will increase further for longer simulation times.

The determined force field parameters (see chapter 5.1) are applicable for all surfactants of the Triton X series, as they differ only in the length of the poly(ethylene oxide) chain. A single 200 ns simulation of TX100 self-assembly has been performed and the progress of the maximum aggregation number is shown in Figure 5.8. As already observed in the TX114 self-assembly simulations, during the first 50 ns of the simulation small aggregates consisting of 30-40 surfactant monomers are formed.

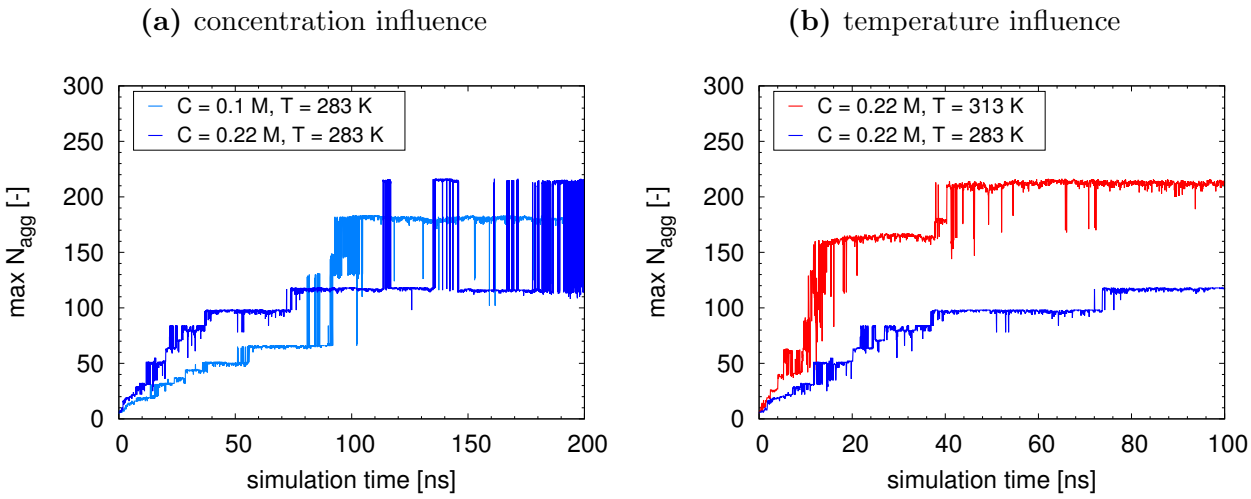

Figure 5.6: (a) Maximum aggregation number of TX114 at two different concentrations. (b) Maximum aggregation number of TX114 at two different temperatures.

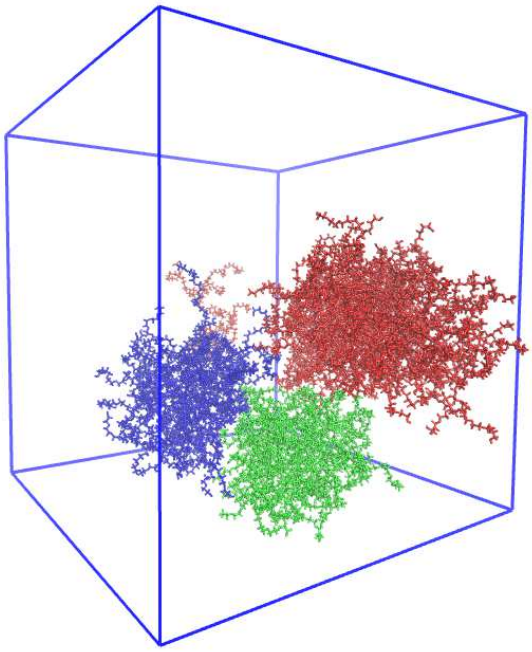

Figure 5.7: A snapshot of TX114 self-assembly simulation at 0.22 mol/L after 100 ns. With different colors are shown three different micelles (aggregation numbers 120, 48, 48). For a clear presentation of the micelles, water and the monomers were removed in this snapshot.

Larger micelles with aggregation numbers around 190 are observed after 60 ns and after 130 ns almost all TX100 monomers aggregate to one micelle with $N_{agg} = 214$. This aggregation number is in agreement with another atomistic MD study of TX100 self-assembly using the GROMOS force field,[178] where all 200 monomers containing in the simulation box converged to a single micelle.[177] Patel et al.[266] reported an experimentally determined TX100 aggregation number $N_{agg} = 287$. Since the experimental aggregation number is larger than the total number of surfactant molecules in the simulation, it is expected that all monomers converge to one micelle. An increase of the system size would probably show even larger micelles.

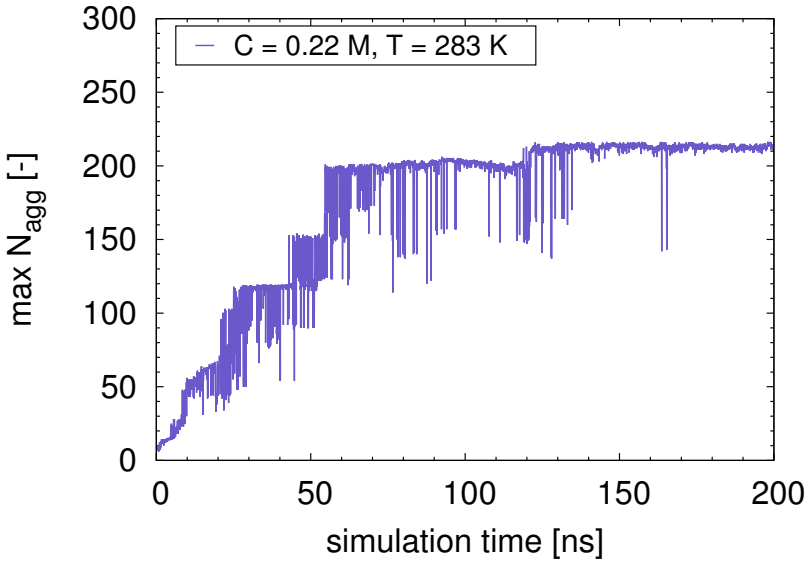

Figure 5.8: Maximum aggregation number of TX100 from a 200 ns self-assembly simulation.

In this chapter, the self-assembly of TX114 and TX100 was simulated with the force field parameters determined in this work (see chapter 5.1). The expected concentration and temperature dependence of the aggregation process could be qualitatively reproduced, which can be interpreted as further validation of the optimized parameters. Aggregation numbers obtained from the simulations are in the same order of magnitude with experimentally determined values. However, it is known that the equilibrium state is difficult to reach with MD simulations of surfactant self-assembly.[181,267] Therefore, a quantitative comparison between simulated and experimentally determined aggregation numbers should be intrepreted with caution.

5.2.3 Prediction of partition equilibria: Influence of micelle size and shape

In this chapter, micelle structures obtained from the self-assembly simulation in chapter 5.2.2 are used to predict partition coefficients of neutral solutes with the model COSMOmic. The influence of the micelle size and shape on the prediction quality with COSMOmic is studied.

To study the influence of the micelle size and shape on the COSMOmic results, different micelles are used for the COSMOmic calculations. The selected TX114 micelle sizes N_{agg}, their sampling (number of averaged micelles per atomic distribution), eccentricity ε, radius of gyration R_{g} and radius of the micelle R_{s} are presented in Table 5.8. These micelle structures are used for the prediction of the partition behavior of ten neutral solutes with COSMOmic. The predicted partition coefficients logK_i between TX114 micelle and water in comparison with experimental data[243] are shown in Table 5.9.

Table 5.8: Structure parameters of averaged TX114 micelles, the angular brackets donate that these values were averaged over the micelles.

N_{agg}	Number of micelles	$< \varepsilon >$	$< R_{\mathrm{g}} >$ [nm]	$< R_{\mathrm{s}} >$ [nm]
32[a]	1673	0.17 ± 0.07	1.72 ± 0.04	2.22 ± 0.05
33[b]	440	0.18 ± 0.06	1.77 ± 0.04	2.28 ± 0.05
65[a]	465	0.33 ± 0.06	2.26 ± 0.06	2.92 ± 0.08
98[b]	1265	0.42 ± 0.03	2.62 ± 0.05	3.38 ± 0.06
116[b]	633	0.64 ± 0.03	3.15 ± 0.10	4.06 ± 0.13
181[a]	511	0.30 ± 0.03	3.26 ± 0.07	4.21 ± 0.09

[a]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.1 mol/L TX114 solution.

[b]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.22 mol/L TX114 solution.

By comparing the RMSE of all logK_i values presented in Table 5.9, no significant differences for the range N_{agg} = 32-98 can be observed, since all overall RMSEs are similar (∼0.40). The predicted partition coefficients slightly overestimate the experimental values, but still provide good prediction quality. The calculated partition coefficients using micelles with aggregation numbers in the range of the experimental values (N_{agg} = 116-181) are in very good agreement with experimental results, since the RMSE = 0.27. When using larger micelles ($N_{\mathrm{agg}} \geq 100$) as input for the COSMOmic calculations, the tendency to overestimate the partition coefficients is reduced. However, usually it is expected that larger micelles are less suitable for COSMOmic as it is more likely that their shape deviate from a sphere.[170] Nevertheless, here larger micelles show a slight increase in prediction quality. A reason could be a different arrangements of surfactants within the micelle. Although the eccentricity for N_{agg} =

181 is only 0.3, the structure is not a sphere where the end atoms of the heads are all located on the surface. By visual inspection of only the hydrophobic region of a micelle of this size it seems that this region has a shape similar to a dumb-bell, probably because two micelles are aggregated to one large micelle. Furthermore, the radial densities show water close to the hydrophobic core. The eccentricity shows still a small value as all atoms are forming a sphere and only by visual inspections and by the radial densities it becomes obvious that this is not a reasonable and suitable micelle structure for COSMOmic. Hence, especially the radial density should also be checked before a micelle is used for COSMOmic.

Table 5.9: Predicted partition coefficients of ten neutral solutes in TX114 micelles at 298 K in comparison with experimental data.[243] All K_i have units of L/kg.

	$\log K_i^{\mathrm{COSMOmic}}$						
Solute	$N_{\mathrm{agg}} = 32^a$	$N_{\mathrm{agg}} = 33^b$	$N_{\mathrm{agg}} = 65^a$	$N_{\mathrm{agg}} = 98^a$	$N_{\mathrm{agg}} = 116^a$	$N_{\mathrm{agg}} = 181^a$	$\log K_i^{\mathrm{exp}}$
Diclofenac	4.01	4.02	4.01	4.07	3.86	3.47	3.82
Ibuprofen	4.26	4.23	4.09	4.10	3.93	3.40	3.73
Ferulic acid	2.71	2.78	2.76	2.77	2.57	2.53	2.20
Hydroxybenzoic acid	1.92	1.96	2.05	2.09	1.90	2.00	1.83
p-Coumaric acid	2.04	2.12	2.22	2.26	2.06	2.17	2.19
Phenol	2.00	2.01	2.00	2.02	1.84	1.79	1.65
Salicylic acid	2.95	2.97	2.95	2.98	2.78	2.64	2.23
Syringic acid	1.78	1.81	1.85	1.85	1.67	1.78	1.64
Vanillic acid	1.87	1.92	2.03	2.06	1.87	1.99	1.88
Vanillin	2.05	2.11	2.08	2.07	1.89	1.89	1.49
RMSE	0.40	0.42	0.40	0.41	0.27	0.27	

[a]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.1 TX114 solution.

[b]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.22 TX114 solution.

The structural parameters (eccentricity factor ε, radius of gyration R_{g} and radius of the micelle R_{s}) of averaged TX100 micelles are given in Table 5.10. Predicted partition coefficients of six neutral solutes between TX100 micelles and water in comparison with experimental results[268,269] are shown in Table 5.11.

Table 5.10: Structure parameters of averaged TX100 micelles, the angular brackets donate that these values were averaged over the micelles.

N_{agg}	Number of micelles	$< \varepsilon >$	$< R_g >$ [nm]	$< R_s >$ [nm]
33[a]	228	0.22 ± 0.10	1.92 ± 0.07	2.48 ± 0.09
34[a]	433	0.21 ± 0.05	1.91 ± 0.04	2.46 ± 0.05
213[a]	357	0.57 ± 0.08	4.24 ± 0.30	5.47 ± 0.40
214[a]	294	0.56 ± 0.08	4.21 ± 0.27	5.43 ± 0.35

[a]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.22 mol/L TX100 solution.

Table 5.11: Predicted partition coefficients of six neutral solutes in TX100 micelles in comparison with experimental data.[268,269] All K_i have units of L/kg.

	$\log K_i^{\text{COSMOmic}}$				
Solute	$N_{agg} = 33$[a]	$N_{agg} = 34$[a]	$N_{agg} = 213$[a]	$N_{agg} = 214$[a]	$\log K_i^{exp}$
3-Methoxyphenol[b]	2.10	2.13	1.73	1.74	1.96
Naphtalene[c]	3.18	3.21	2.37	2.38	3.10
Phenanthrene[c]	4.01	4.03	3.07	3.09	4.16
Phenol[b]	1.95	1.98	1.58	1.59	1.79
Pyrene[c]	4.31	4.33	3.31	3.33	4.49
Vanillin[b]	2.00	2.01	1.58	1.59	1.90
RMSE	0.14	0.15	0.74	0.73	

[a]The micelle structures used for COSMOmic calculations were taken from a 200 ns MD simulation of 0.22 TX100 solution.
[b]The partition coefficients are measured and calculated at 293 K.[268]
[c]The partition coefficients are measured and calculated at 298 K.[269]

The calculated partition coefficients using small TX100 micelle structures ($N_{agg} = 33$ and $N_{agg} = 34$) are in excellent agreement with experimental data (RMSE = 0.14 and 0.15, respectively). The usage of small TX100 micelles (aggregation number ∼33) shows better prediction quality (RMSE ∼0.15) compared to the calculations using TX114 micelles of the same size (RMSE ∼0.40). Since no significant difference in the micelle structures are observed (aggregation number, micelle radius and eccentricity factor are in the same range), it can be concluded that the deviation is due to the different solutes which were studied. The predicted partition coefficients using larger micelles ($N_{agg} = 213$ and $N_{agg} = 214$) underestimate the experimentally determined partition coefficients with RMSE = 0.74 and RMSE = 0.73, respectively. Consequently, when using large micelles ($R_s > 5$ nm) outliers can occur. The reason is most probably the deviation of the structure from a sphere (see ε in Table 5.10). This usually leads to deviations in the predictions, as in COSMOmic the micelle structure

is assumed as a perfect sphere surrounded by a water shell.

The main factor that influences the quality of the COSMOmic predictions is the form of the micelle, defined by the eccentricity factor. Therefore, in order to achieve appropriate prediction quality with COSMOmic, the micelles formed during MD simulations, need to be analyzed and selected according to their structure. Although cylindrical micelles ($\varepsilon \geq 0.5$) can also provide good prediction quality (see Table 5.9), outliers are more likely to occur. On the other hand, when using small spherical micelles ($R_s < 4$ nm , $\varepsilon \leq 0.5$) as input structures for the COSMOmic calculation, all predicted partition coefficients are in very good agreement with experimental data (RMSE ≤ 0.42). Therefore, it can be recommended to select micelle structures from MD according to their structural parameters (radius of the micelle and eccentricity factor) and by checking the radial densities.

5.3 Summary

It has been demonstrated that using the force field parameters for nonionic surfactants from the Trition X series introduced in this work, it is possible to simulate the self-assembly of TX114 and TX100 and to reproduce qualitatively the concentration and temperature dependence of the micelles formation. The aggregation number increases with increasing temperature and surfactant concentration, which is in agreement with experimental data.[265] Micelle structures taken from the self-assembly simulations were used for predictions of partition equilibria with COSMOmic. In order to achieve statistical reliable results, averaged atomic distributions for the micelle structure were used, which reduces the effect of outliers. The predictions are in very good agreement with experimental results for neutral molecules. Moreover, the influence of micelle size and shape on the prediction quality was analyzed. Outliers are more likely to occur when using large and more cylindrical micelles, whereas reproducible results were obtained when using small spherical micelles as an input for COSMOmic.

In the next chapter COSMOmic calculations are further validated by comparing free energy profiles.

6 Free energy profiles and partition equilibria in various micelles: Comparison between MD simulations, COSMOmic and experiments

In this chapter, the partition behavior of solutes in all types of micelles (non-ionic, zwitterionic, anionic and cationic) is investigated, whereby COSMOmic, MD simulations and experiments are compared. To gain insights on the partition behavior as a function of the position within the micelle, free energy profiles are obtained and analyzed. Furthermore, the applicability of MD simulations (umbrella sampling method) and COSMOmic to predict partition behavior of charged solutes is evaluated. To make the quantitative comparison possible, partition coefficients of dissociated molecules in cationic micelles are determined with micellar liquid chromatography. The results of the following sections were partly published as Yordanova, D.; Ritter, E.; Gerlach, T.; Jensen, J.H.; Smirnova, I.; Jakobtorweihen, S. **Solute Partitioning in Micelles: Combining Molecular Dynamics Simulations, COSMOmic and Experiments** in Journal of Physical Chemistry B.[270] Some results are based on the supervised master's thesis of Jan-Hendrik Jensen.[106]

6.1 Micellar liquid chromatography for CTAB micelles

The MLC method has been established as robust and reliable technique to obtain partition coefficients of neutral solutes in micellar systems.[271] However, the determination of dissociated compounds is more challenging. Only a few partition coefficients of ionizable molecules are available in the literature.[42,105] One aim of this work is to evaluate MD simulations with the US method and the COSMOmic model for the calculation of partition coefficients in micelles with special focus on charged solutes. In MD simulations one of the main limitations is the force field. Unfortunately, no CHARMM force field parameters (the force field used for the surfactants as described before) are available for ionizable molecules, for which experimental partition coefficients are available. To make the quantitative comparison possible, solutes with available parameters in the CHARMM General Force Field (CGenFF)[112] were selected and their partition coefficients are determined with MLC. Therefore, partition coefficients of 4-hydroxybenzoic acid, 4-methoxyphenol and phloretic acid in CTAB micelles were measured at three pH values (2, 7 and 10.5) to consider different

dissociation states of the molecules. The structural formulas of the selected solutes in their neutral form are shown in Figure 6.1. At pH = 2 all solutes will be protonated (no overall charge). The dissociation degrees of the compounds at pH = 7 and pH = 10.5 in water are shown in Table 6.1. The determined partition coefficients are listed in Table 6.2. Since 4-methoxyphenol is 99.89 % protonated at pH = 7, nearly identical partition coefficients at pH = 2 and pH = 7 are expected. The deviation of 0.18 log units can be related to inaccuracies of the experiment.

Table 6.1: Degrees of dissociation in water at $T = 298$ K. α_0 corresponds to the undissociated molecules, α_1 and α_2 are the degree of dissociation of the first (carboxyl group, in case of 4-methoxyphenol hydroxyl group) and second dissociable functional group (hydroxyl group), respectively.

	pH = 7			pH = 10.5		
Solute	α_0 [%]	α_1 [%]	α_2 [%]	α_0 [%]	α_1 [%]	α_2 [%]
4-Hydroxybenzoic acid	0.24	99.55	0.21	0.00	12.89	87.11
4-Methoxyphenol	99.89	0.11	-	21.59	78.11	-
Phloretic acid	0.16	99.53	0.31	0.00	9.28	90.72

Table 6.2: Experimentally determined partition coefficients of three ionizable solutes in CTAB micelles. The standard deviation was derived from three measurements. All K_i have units of L/kg.

	$\log K_i^{\mathrm{exp}}$		
Solute	pH = 2	pH = 7	pH = 10.5
4-Hydroxybenzoic acid	2.64 ± 0.04	3.91 ± 0.09	-
4-Methoxyphenol	2.38 ± 0.01	2.20 ± 0.01	3.03 ± 0.03
Phloretic acid	2.94 ± 0.21	3.99 ± 0.68	-

As expected, the partition coefficients increase with increasing pH value. This effect can be attributed to the strong electrostatic interactions between the deprotonated groups (negatively charged) and the cationic CTAB head groups. Similar effects were observed in previous studies.[42,105] According to the protolytic equilibrium at pH = 10.5 the second dissociable functional group (hydroxyl group) of 4-hydroxybenzoic acid and phloretic acid is also dissociated (see Figure 6.1 for the molecule structures). Due to increasing net charge of the solutes at higher pH, increasing attraction forces between the cationic groups of CTAB and the dissociated molecules are expected. However, the calculated pK_a values refer to an aqueous solution and it is known that the description of

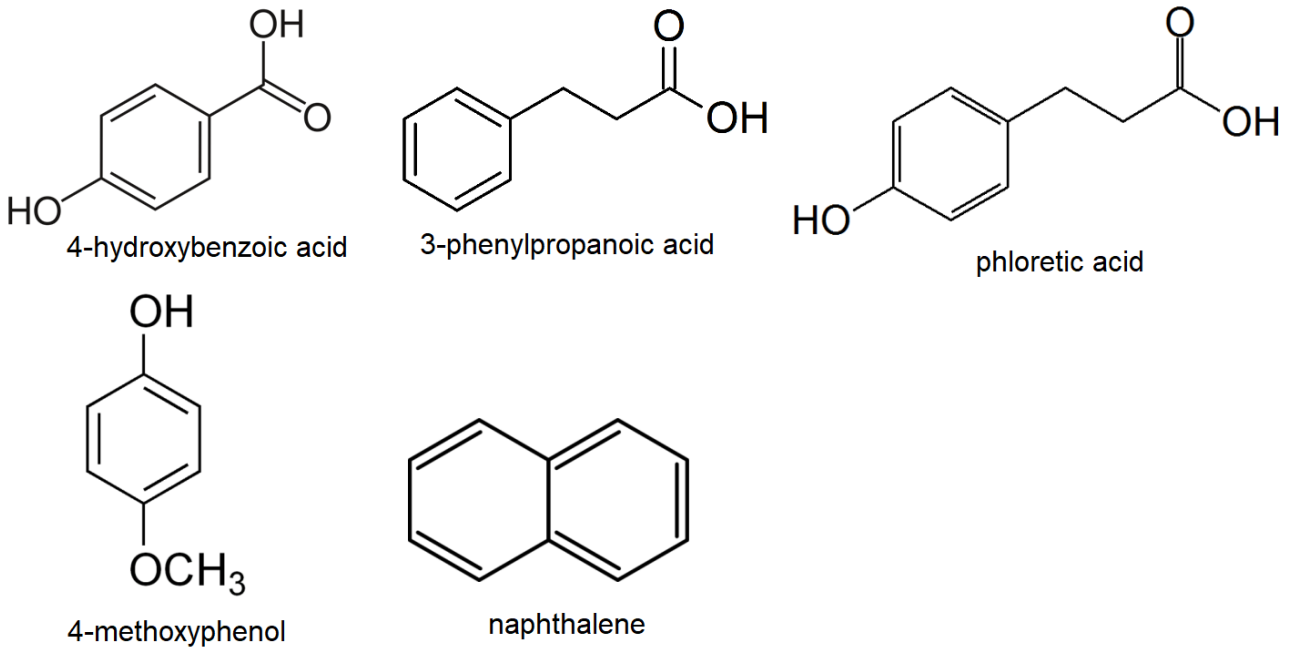

Figure 6.1: Structural formulas of the selected solutes in their neutral form.

the protolytic equilibrium can differ significantly in membranes.[226] For example, Beschiaschvili and Seelig[227] showed shifted pK_a values of peptides in contact with lipid membranes. Schreier et al.[226] demonstrated that the apparent pK_a of the drug tetracaine decreases in the presence of membranes, being a function of membrane concentration. The change of the pK_a from the aqueous phase to the membrane can be estimated theoretically based on transfer free energies obtained from MD simulations.[194]. Shifts of ionization constants are observed in micellar systems as well.[221,228] Rodgers and Khaledi[228] reported significantly increasing ionization constants of amino acids in micellar solutions. However, it is known that electrostatic interactions are the driving forces causing changes in the pK_a at the membrane/water interface.[218] Thus, in the case of cationic CTAB micelles and dissociated acids, lower pK_a values on the micellar surface than in water are expected. A similar behavior was demonstrated for the binding of valsartan dianion to CTAB micelles.[221] Therefore, increasing partition coefficients of 4-hydroxybenzoic acid and phloretic acid at pH = 10.5 are expected. However, the determined partition coefficients at pH = 7 and pH = 10.5 are nearly identical (the values at pH = 10.5 are not shown). It has been demonstrated that the MLC method is very sensitive to the micelle-solute interactions.[42] Due to very strong binding between the micelle and the solute (for example due to strong electrostatic interactions in case of cationic surfactant and dissociated solutes) the evaluation with the retention models is limited.[42,105] This behavior is designated as overbinding. Čudina et al.[221] demonstrated that binding constants, obtained by using MLC and UV spectrophotometry show

great discrepancy. Hence, in the case of 4-hydroxybenzoic acid and phloretic acid MLC is probably not sensitive enough to estimate the protolytic equilibrium of the second functional group. As the determined partition coefficients in this case cannot be considered as reliable, the values are not included in Table 6.2. The determination of a broader range of partition coefficients using the MLC method could be challenging and time consuming. Therefore, other technique, such as the MEUF method could be more suitable for dissociated compounds.

6.2 Simulation details

The MD simulations were performed with GROMACS package version 5.1.1[272] or version 4.6.7[249] using the CHARMM36 force field[129,273] and the CHARMM TIP3P water model.[122] TX114 was modeled with the force field parameters optimized in this work (see chapter 5.1). Parameters for the bromide ions are taken from Joung and Cheatham[274] who optimized them for TIP3P water. In all simulations the temperature was fixed at 298 K using the Nosé-Hoover thermostat[246] (τ_t = 1 ps). The time step for all simulations was 2 fs. The Lennard-Jones interactions were cut off at 1.2 nm and switched from 1.0 nm on (0.8 nm for GROMACS version 4.6.7), where the force was switched for GROMACS version 5.1.1 (and both the potential and the force for GROMACS version 4.6.7). For the Coulomb interactions, a real space cutoff of 1.2 nm was applied (1.0 nm for GROMACS version 4.6.7). The Particle Mesh Ewald method was used for the long-ranged electrostatic interactions.[133,250] The bonds to hydrogens were constrained with the LINCS algorithm.[251]

6.2.1 Preassembled micelles

In this chapter, preassembled micelles were used as starting structures. They were generated with PACKMOL,[275] containing 40 surfactant monomers. To obtain spherical starting structures, surfactant atoms were restricted in two defined spheres. Water molecules were added such that a concentration of 0.08 M was reached. In case of ionic surfactants, counterions were added. After an energy minimization with the steepest descent method, a simulation in the NPT ensemble with the Berendsen barostat[247] (isotropic, coupling constant τ_p = 5 ps) was performed for 200 ps in the case of SDS and CTAB micelles and 10 ns for TX114 and HePC, respectively. Followed by 40 ns NPT simulations with the Parrinello-Rahman barostat[248] (p = 1 bar, isotropic coupling constant τ_p = 5 ps).

6.2.2 Umbrella sampling protocol

The umbrella sampling method[152,276,277] was applied to calculate the free energy profiles of four neutral molecules in different micelles: 4-hydroxy-benzoic acid in TX114 and CTAB micelles, 3-phenylpropanoic acid in HePC, 4-methoxyphenol and naphthalene in SDS. The structures of the solutes and of the surfactants are shown in Figure 6.1 and Figure 4.1, respectively. Furthermore, 3-phenylpropanoic acid and 4-hydroxybenzoic acid were investigated in their dissociated states in HePC and CTAB micelles, respectively. All simulated systems are listed in Table 6.3. Force field parameters for the solutes are taken from the CGenFF[112] and were assigned using the ParamChem tool,[144,145] where the highest charge penalty = 3.845 and the highest parameter penalty = 6.500 were obtained for 3-phenylpropanoic acid (neutral) and phloretic acid (neutral and first dissociation equilibrium). All parameters for 4-hydroxybenzoic acid, 4-methoxy-phenol and naphthalene have penalties equal to zero.

The simulated preassembled micelles (see chapter 6.2.1) were used as starting structures for the umbrella sampling (US) simulations. To obtain better statistics and error estimates, two solute molecules were added to the simulated TX114, HePC and SDS micelles, described in chapter 6.2.1. The molecules were placed into the water phase, the overlapping water molecules were removed. In case of dissociated solutes, counterions were added explicitly to keep the total charge of the system neutral. Subsequently, NPT simulations were performed for 10-40 ns using the Parrinello-Rahman barostat[248] ($p = 1$ bar, coupling constant $\tau_p = 5$ ps). In case of CTAB, a system containing a preassembled micelle, two solute molecules and counterions was directly prepared with PACKMOL. After an energy minimization, a simulation in the NPT ensemble with Berendsen barostat[247] was performed for 200 ps, followed by a 40 ns NPT simulation with the Parinello-Rahman barostat.[248] To obtain starting configurations for the US simulations, the solutes were pulled (pull rate = 0.002 nm/ps, force constant = 3000 kJ/mol nm^2) towards and away from the COM of the micelle. In case of charged solutes, the micelles did not remain stable after the pulling simulation and two smaller aggregates were observed. This is due to repulsive forces between the charged solute and the hydrophobic core of the micelle. As the partitioning of the charged solute in the micelle core is unfavorable, the micelle falls apart such that the solute makes contact with water and is located between the two aggregates at the COM of all surfactants. To overcome this effect, restraints were employed to limit the movement of surfactant monomers. Two atoms from the hydrophobic part (first and terminal carbon) of HePC and CTAB molecules were restricted by applying a flat bottom potential with a force constant = 10000 kJ/mol and a flat bottom radius = 0.4 nm (see chap-

ter 2.3.2). Finally, 36 umbrella sampling simulations were performed for each solute, where the distance between each window was 0.1 nm. The force constant for the umbrella potential was 3000 kJ/mol. By using two solutes two free energy profiles were obtained and an averaged profile with error bars was calculated.[235] The performed US simulations are listed in Table 6.3.

Table 6.3: Simulation details of the umbrella sampling simulations.

Solute	Solute charge	Surfactant	Simulation time
4-Hydroxybenzoic acid (HBA)	0	TX114	36 x 40 ns
3-Phenylpropanoic acid (PPA)	0	HePC	36 x 40 ns
3-Phenylpropanoic acid (PPD)	-1	HePC	36 x 40 ns
4-Methoxyphenol (MTP)	0	SDS	36 x 40 ns
Naphthalene (NAPH)	0	SDS	36 x 40 ns
4-Hydroxybenzoic acid (HBA)	0	CTAB	36 x 40 ns
4-Hydroxybenzoic acid (HBD1)	-1	CTAB	36 x 40 ns
4-Hydroxybenzoic acid (HBD2)	-2	CTAB	36 x 40 ns

To calculate the partition coefficient between the micelle and the aqueous phase, it is crucial to define a border between the micellar phase and the aqueous phase. The free energy profiles, obtained with the US method were used to distinguish between the micelle and the water phase: the water phase begins where the free energy has reached a value near zero (maximal deviation < 0.40 kJ/mol). The chosen criteria are 2.6 nm for HePC micelles, 2.5 nm for CTAB and SDS micelle and 3.2 nm for TX114, respectively.

6.2.3 COSMO-RS and COSMOmic

Surfactant conformers were selected from micelles with respect to their solvent accessible surface.[240] The required surfactant DFT/COSMO calculations were performed with a single point calculation with Turbomole 6.6.[278] Conformers for the solute molecules were created with COSMOconf version 3.0[279] and the corresponding DFT/COSMO calculations were performed with Turbomole 6.6.[278] The COSMOtherm[280,281] (version C3.0 Release 15.01) implementation of COSMO-RS and COSMOmic with the BP_TZVP_C30_1501 parametrization was used. The same DFT/COSMO calculations and COSMOtherm version were also used in chapter 7.

6.3 Calculated free energy profiles and partition coefficients

The US method and the COSMOmic model are applied to predict partition equilibria of various solutes in micelles. As COSMOmic was introduced for neutral molecules, predictions for charged solutes should be interpreted with caution. Therefore, a comparison between MD and COSMOmic free energy profiles in micelles is beneficial. By comparison to experimental partition coefficients, the applicability of both models to predict the partition behavior of solutes in micelles is evaluated.

For the calculation of partition coefficients from free energy profiles, different approaches are possible (see chapters 2.3.3 and 2.4.2). In MD studies mostly equation 2.61 is used, where a border between the micelle and the water phase is defined. However, the original COSMOtherm implementation uses a different equation.[52] To obtain an equation that is stable also for hydrophilic solutes, Klamt et al.[52] introduced the concept of a water correction (equation 2.74, chapter 2.4.2). In this chapter, both definitions are applied and compared for the calculations with COSMOmic. Partition coefficients calculated with the COSMOtherm implementation of COSMOmic are designated as $\log K_i^{\mathrm{COSMOmic,Wcorr}}$. To be consistent with the MD results, partition coefficients from COSMOmic free energy profiles were also calculated using equation 2.61 (designated as $\log K_i^{\mathrm{COSMOmic,Pdef}}$). It is important to note that for equations 2.61 and 2.74 the volume corrected profile should be used (see Figure 4.3, chapter 4.4). Whereas, for unmodified profiles the volumes of the layers should not be used in equations 2.61 and 2.74, as the volume contribution is already incorporated in the profile.

The predicted and experimental partition coefficients are summarized in Table 6.4. The experimental error scales of the partition coefficients (see Table 6.2) are mean standard deviations of three determinations. However, partition coefficients obtained with different experimental techniques can vary widely and an experimental error of 0.3 log units can be assumed.[201,282]

In reference 161 the influence of the applied equation on the membrane/water partition coefficients was investigated. Deviations within ∼0.3 log units are observed between the two equations. In the case of micelles higher deviations (∼0.5 log units) than in membranes are observed, which can be attributed to the different geometries. Since micelles are more flexible than lipid bilayers, more water molecules are allowed to be located inside the micelle phase. The water content inside the simulated zwitterionic micelle phase is 85 water molecules per surfactant, which is much higher compared to the water to lipid ratio in a lipid bilayer phase (17 water molecules per DMPC lipid). As expected, higher

Table 6.4: Predicted partition coefficients in comparison with experimental data, all K_i have units of L/kg. All predictions were obtained with micelles consisting of 40 surfactants.

Solute	Surfactant	$\log K_i^{\mathrm{US,Pdef}}$	$\log K_i^{\mathrm{COSMOmic,Pdef}}$	$\log K_i^{\mathrm{COSMOmic,Wcorr}}$	$\log K_i^{\mathrm{exp}}$
HBA	TX114	2.03 ± 0.16	1.55	2.13	1.83[23]
PPA	HePC	2.63 ± 0.23	1.88	2.33	-
PPD	HePC	1.06 ± 0.32	0.61	1.09	-
MTP	SDS	2.17 ± 0.10	1.57	2.14	1.62[283]; 1.73[284]
NAPH	SDS	2.73 ± 0.08	2.75	3.33	3.00[285]; 3.07[283]
HBA	CTAB	2.51 ± 0.03	2.04	2.60	2.64
HBD1	CTAB	1.69 ± 0.07	2.97	3.46	3.91
HBD2	CTAB	1.32 ± 0.02	4.03	4.45	-
PLA	CTAB	-	2.34	2.90	2.94
PLD1	CTAB	-	3.37	3.85	3.99
PLD2	CTAB	-	4.61	5.02	-
MTP	CTAB	-	2.29	2.85	2.38; 2.20
MTD	CTAB	-	3.63	4.11	3.03

partition coefficients are obtained when using the water correction, than those calculated with the phase boundary definition. Significant deviations can occur when using equation 2.61 for hydrophilic compounds. In this case the concept of a water correction has to be applied. Below, the calculated free energy profiles and partition coefficients are discussed in detail according to the surfactant type.

6.3.1 Nonionic surfactants

For the modeling of partition equilibria in nonionic micelles, TX114 was chosen as representative nonionic surfactant. Figure 6.2 shows free energy profiles of 4-hydroxybenzoic acid (HBA) in its neutral state in a TX114 micelle. The hydrophilic head group (oxyethylene chain) of TX114 is larger than its hydrophobic part, depicted also by the radial distribution of the head group. The free energy minimum is located completely in the hydrophilic part of the micelle (palisade layer) at a distance of 1.35 nm (COSMOmic profile) from the center of the micelle. The favorable partitioning of HBA inside the palisade layer can be attributed to the polar groups of the molecule. The free energy profiles from the US simulations and the COSMOmic calculation are in good agreement, both showing the same location and depth of the free energy minimum. The predicted partition coefficients are in good agreement with the experimental value, as the deviations are lower than 0.3 log units (see Table 6.4).

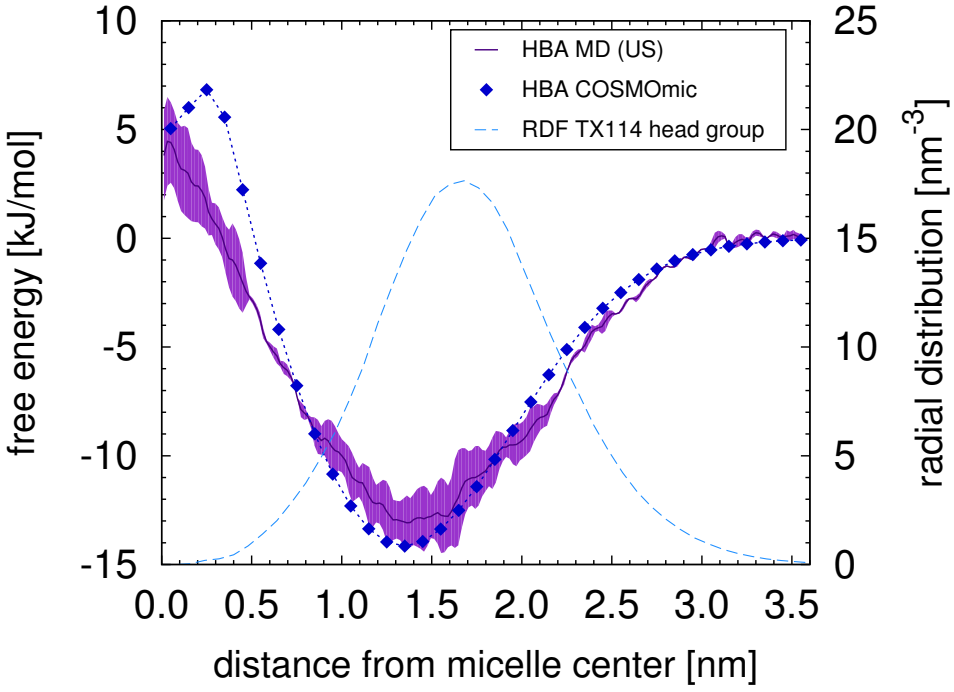

Figure 6.2: Free energy profiles of 4-hydroxybenzoic acid (HBA) in a TX114 micelle calculated from MD simulations (US method) and with the COSMOmic model. The colored areas show the error margin of the MD simulation results. Radial distribution function (RDF) of the surfactant head group atoms (dashed line) calculated from MD simulation without HBA is plotted on the secondary y-axis.

6.3.2 Zwitterionic surfactants

The zwitterionic detergent HePC (miltefosine) is often used as model for biomembranes due to its structural similarity to the phospholipids present in biomembranes and the easier preparation of micelles compared to liposomes.[26–28,286] Therefore, HePC was chosen as an example for a zwitterionic surfactant. Free energy profiles of 3-phenylpropanoic acid in the neutral (PPA) and dissociated (PPD) state are shown in Figure 6.3.

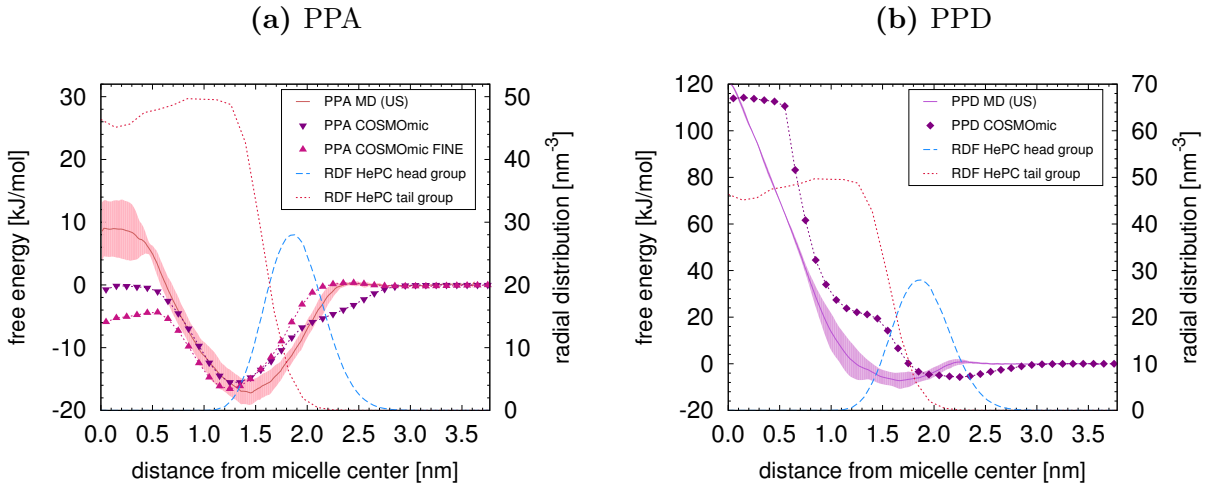

Figure 6.3: Free energy profiles of 3-phenylpropanoic acid in neutral (PPA) and dissociated (PPD) state in a HePC micelle calculated from MD simulations (US method) and with COSMOmic. The colored areas show the error margin of the MD simulation results. Radial distribution functions of the surfactant head and tail group atoms calculated from MD simulation without solutes are plotted on the secondary y-axes.

The free energy profiles for PPA from US simulations and from COSMOmic are similar. However, differences are obtained in the interfacial region. It is well known that COSMO-RS shows systematic errors for secondary and tertiary amines, which is likely due to the description of the hydrogen bonds of the amine group within the model.[50,163] In contrast to this, the phosphadityl-choline (PC) head group of HePC contains a quaternary ammonium group which is not capable of accepting hydrogen bonds. Nevertheless, similar deviations in the interfacial region were observed for lipid bilayers containing the PC head group,[161,287] and it was shown, by comparing lipids with other head groups, that this deviation is due to the PC head group.[161] In the classic COSMO-RS model the electrostatic misfit energy is calculated locally for surface segments in close contact. Thus, any nonlocal, long range electrostatic interactions are neglected[288] but might be of importance for the charged quaternary ammonium group. Recently, a new COSMO-RS parametrization called BP-TZVPD-FINE was introduced,[288] including improvements for the descrip-

tion of hydrogen-bonding and electrostatic interactions. The free energy profile of PPA was calculated again using the BP-TZVPD-FINE parametrization and is shown in Figure 6.3a. It can be seen, that by using the BP-TZVPD-FINE parametrization the COSMOmic result is in better agreement with the MD result in the interfacial region. Further deviations between the methods are observed in the hydrophobic core of the micelle. The free energy profile from US simulations shows a free energy of ∼9 kJ/mol in the hydrophobic core, which can be associated to the polar carboxyl group of the acid. In contract, COSMOmic predicts ∼0 kJ/mol free energy in the hydrophobic core and ∼ −6 kJ/mol when using the FINE parametrization. It is assumed that the free energy in the hydrophobic center of biomembranes can be modeled by cyclohexane/water transfer free energies.[235] Unfortunately, for 3-phenylpropanoic acid no experimental cyclohexane/water transfer free energy was found in the literature. Nevertheless, Abraham et al.[289] obtained a free energy of 4.85 kJ/mol for benzoic acid and 13.69 kJ/mol for propanoic acid, respectively. It is expected that the transfer free energy of 3-phenylpropanoic acid is between those of benzoic and propanoic acid, which is in agreement with the US method prediction of ∼9 kJ/mol. Therefore, it can be assumed that the COSMOmic calculations underestimate the free energy of PPA in the hydrophobic core of the HePC micelle. The free energy minimum of PPA is in agreement for both models, which is located between the PC head group and the hydrophobic region of HePC. For the dissociated form of 3-phenylpropanoic acid (PPD) large energy barriers in the hydrophobic core are obtained (see Figure 6.3b), as expected. The COSMOmic profile shows a small change in the slope between the hydrophilic and hydrophobic groups, which can be associated with overpredicted repulsion forces when the molecule is passing through regions with different hydrophobicity. Both methods show deviations in the location of the global minimum, which is predicted at ∼1.65 nm by the US simulations and at ∼2.25 nm by COSMOmic, respectively. Unfortunately, experimentally determined partition coefficients for 3-phenylpropanoic acid in HePC were not found in the literature. Therefore, the partition behavior prediction can be evaluated only qualitatively. Miyazaki et al.[218] demonstrated, that due to interfacial pK_a shift protonated fatty acids have higher partition coefficients in dimyristoylphosphatidylcholine (DMPC) bilayer compared to their dissociated form. Furthermore, partition coefficients of 5-phenylvaleric acid in liposomes were found in the literature,[35,37,290] which has only two methylene groups more than PPA. A partition coefficient of 2.94 log units and 3.17 log units was obtained for protonated 5-phenylvaleric acid in DMPC and DOPC, respectively.[290] A value of 0.50 log was determined for the charged form in DMPC.[35] Correspondingly, the calculations in this work show decreasing partition coefficients for dissoci-

ated 3-phenylpropanoic acid compared to the neutral solute. As the chemical structures of 5-phenylvaleric acid and 3-phenylpropanoic acid are similar and also the structure of HePC with DMPC and DOPC, the predictions are most likely qualitatively correct, for a quantitative assessment experimental data is needed.

6.3.3 Anionic surfactants

In this chapter, the partitioning between SDS micelles and water of two molecules with different hydrophobicity was investigated. Free energy profiles of 4-methoxyphenol (MTP) and naphthalene (NAPH) are shown in Figure 6.4.

(a) MTP **(b)** NAPH

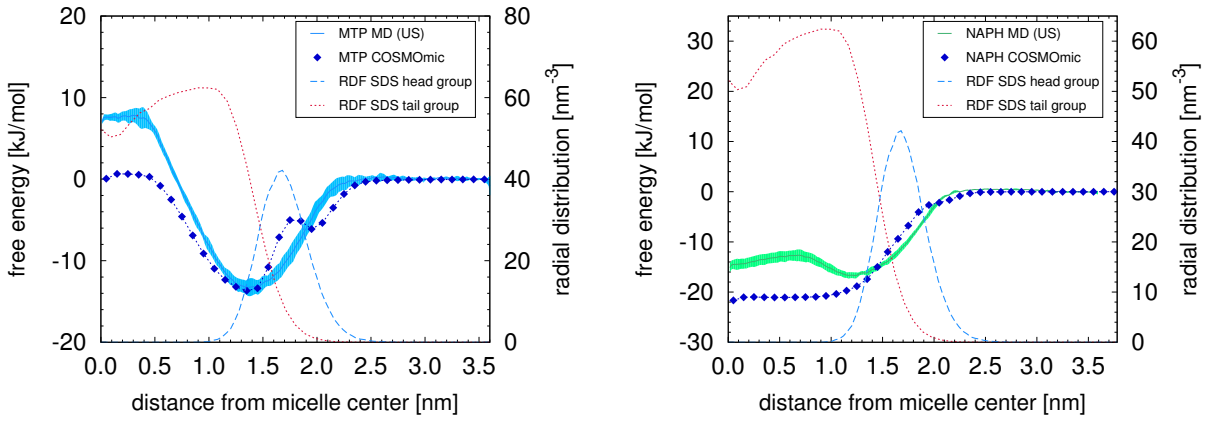

Figure 6.4: Free energy profiles of 4-methoxyphenol (MTP) and naphthalene (NAPH) in a SDS micelle calculated from MD simulations (US method) and with COSMOmic. The colored areas show the error margin of the MD simulation results. Radial distribution functions of the surfactant head and tail group atoms calculated from MD simulation without solutes are plotted on the secondary y-axes.

For both solutes, the free energy profiles calculated with the US method and with COSMOmic show some differences. For MTP deviations at similar positions as for PPA and PPD in HePC micelle are observed. In the hydrophobic core of the micelle the US profile shows a free energy of ∼7.5 kJ/mol, whereas the COSMOmic calculation results in 0 kJ/mol. Rich and Harper[291] obtained transfer free energy of 5.61 kJ/mol (at 296 K) for MTP in cyclohexane/water. Hence, the prediction of the MD simulations can be considered as reasonable, whereas COSMOmic underestimates the free energy in the micelle core. COSMOmic predicts a change in the slope of the profile in the region between the hydrophilic and hydrophobic part of the micelle, which is not expected. Similar behavior was predicted for PPD in HePC micelle. In both cases the changes occur, when the molecules are located in the transition re-

gion between the hydrophilic head groups and the hydrophobic carbon chain (see Figure 6.5). The deviation can be attributed to overpredicted repulsion in the transition region. The partition coefficient of MTP in SDS calculated with COSMOmic using equation 2.61 is in the best agreement with experimental data (absolute deviation 0.09 log units). The predictions from the MD simulations and the COSMOtherm implementation of COSMOmic lead to identical values, which overestimate the experimental partition coefficients with maximal absolute deviation of 0.53 log units (see Table 6.4).

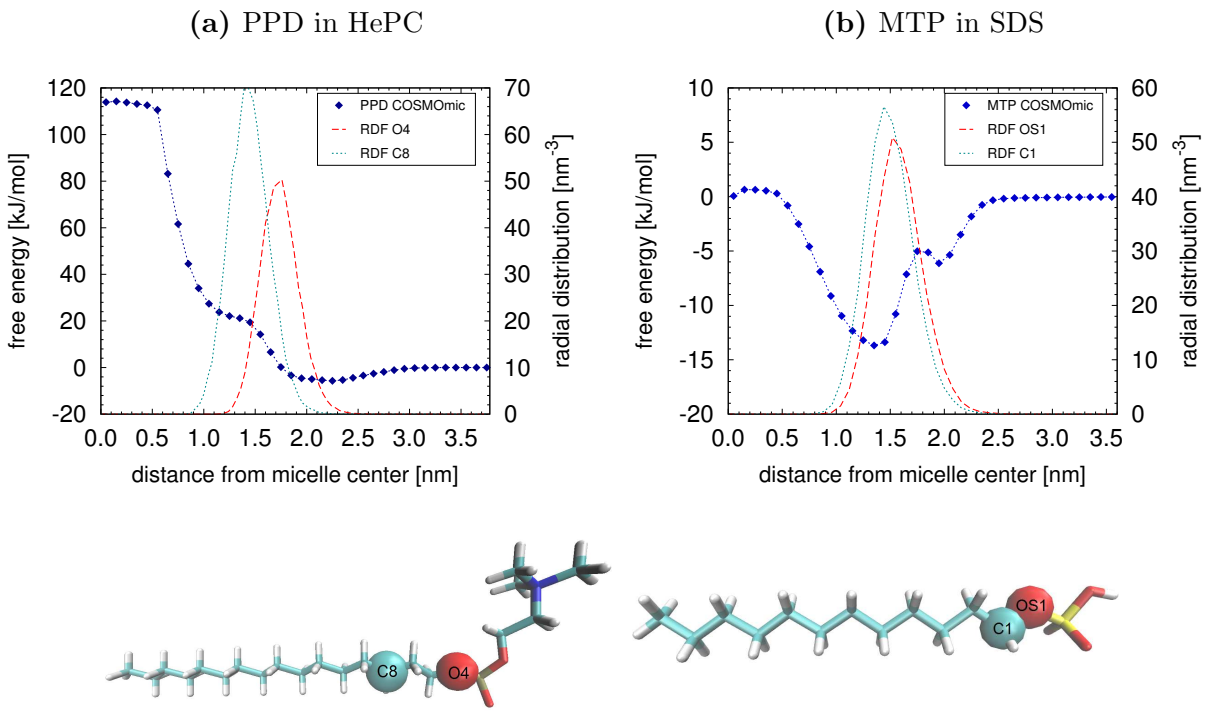

Figure 6.5: (top) Free energy profiles calculated with COSMOmic: PPD in a HePC micelle (a) and MTP in a SDS micelle (b). Radial distribution functions of the highlighted atoms are shown. (bottom) Structures of the corresponding surfactants.

In order to investigate partition behavior of a hydrophobic compound, free energy profiles of naphthalene in SDS are determined. As expected, the solute is preferably located in the hydrophobic region of the micelle (see Figure 6.4b). According to the profile calculated with the US method, a local minimum is observed between the hydrophilic and hydrophobic groups, whereas COSMOmic predicts a constant free energy within the whole hydrophobic region. The free energy at the minimum is ~ -16.5 kJ/mol according to the US profile and a value of ~ -21.5 kJ/mol is predicted by COSMOmic. Abraham et al.[289] obtained -19.29 kJ/mol free energy of transfer for naphthalene in alkane/water. Thus, both methods result in reasonable free energies of NAPH in the hydrophobic micelle core. Therefore, the predicted partition coefficients are also in good agreement with experimental results (see Table 6.4). Although

both free energy profiles differ in the hydrophobic region (0 to 1.5 nm from micelle center), the obtained partition coefficients are similar. The deviation to experimental values is in the range of 0.3 log units.

6.3.4 Cationic surfactants

In order to investigate partition equilibria of ionizable compounds in cationic micelles, 4-hydroxybenzoic acid as solute and CTAB as widely used cationic surfactant were selected. As described before, the partition coefficients in CTAB micelles are higher than those in nonionic micelles because of stronger electrostatic interactions.[42,60,105] Furthermore, the partition coefficients of the dissociated solutes increase significantly compared to those of the protonated form.[42] To evaluate the applicability of MD simulations and COSMOmic to predict these effects, 4-hydroxybenzoic acid was investigated in neutral and dissociated states (see also Table 6.3). The obtained free energy profiles are shown in Figure 6.6. Please note that a micelle with aggregation number of 40 is used for both the MD simulations and COSMOmic calculations. The influence of the micelle size on the free energy profiles and partition coefficients of charged compounds will be discussed in detail below. In the case of HBA, the free energy profiles obtained with the US method and calculated with COSMOmic are in good agreement regarding the shape of the profile and the depth of the minimum. COSMOmic predicts the minimum at 1.75 nm from the COM of the micelle. In the same region the maximum of the radial distribution of the hydrophilic group is obtained. The free energy minimum in the US profile is slightly shifted to the hydrophobic group. In the hydrophobic core a free energy of 13 $\pm$ 2 kJ/mol is predicted by the US method and 19 kJ/mol by COSMOmic, respectively. Both models overestimate the experimental cyclohexane/water transfer free energy of 10.09 kJ/mol.[292] Still, the free energy obtained from MD simulations is in close agreement with the experimental data when considering the error of $\sim$2 kJ/mol in the hydrophobic core. The predicted partition coefficients of HBA in CTAB micelle are higher than in TX114 (see Table 6.4), as expected. Good agreement with experimental results can be achieved by using the US method. The COSMOmic prediction with equation 2.61 leads to a higher deviation of 0.60 log units, only by using the water correction (equation 2.74) COSMOmic is in agreement with the experimental result. Significant deviations are obtained for the charged solutes. Both methods predict huge energy barriers in the hydrophobic core of the micelle, but show crucial differences at the free energy minimum (see Figure 6.6). According to the US profiles, no decrease in the free energy minimum for the dissociated solutes is obtained compared to the neutral form; the free energy at the minimum of HBA and HBD1 re-

mains relatively constant (~ -10 kJ/mol), which does not correspond to the experimental findings (see Table 6.4). Thus, enhanced electrostatic interactions between the cationic head group of CTAB and the dissociated acid are not observed in the MD simulations. This could be attributed to the electrostatic model in the force field and the force field parameters. Force field parameters for CTAB are taken from the CHARMM force field for lipids. Due to different optimization procedures the parameters in different force fields can differ significantly. This is also the case for the charge distribution of the phosphocholine group. Surprisingly, in the CHARMM FF the nitrogen is negatively charged, the positive charge of the head group is distributed between the hydrogens. In contrast, in other all-atom force fields (e.g., OPLS AA[128], Slipids[293] and Amber Lipids FF[125]) the nitrogen is positively charged. For a better understanding of the force field parameters influence, a comparison with other all-atom force fields is reasonable. However, the solutes of interest have to be parametrized within the same force field, as the FF parameters are not transferable between different force fields. Moreover, to reproduce processes driven by electrostatic interactions correctly, the usage of polarizable force field[294,295] would be the most appropriate choice. In contrast to the MD results, different free energies at the minimum are obtained by the COSMOmic calculations. The calculated minimum free energy is -15 kJ/mol for HBA, -20 kJ/mol and -27 kJ/mol for HBD1 and HBD2, respectively. Hence, an increase in the partition coefficient is predicted for the dissociated states. The calculated partition coefficient of HBD2 is 1 log unit higher than that of HBD1, which corresponds to the expected behavior. Therefore, it can be assumed that the COSMOmic model is able to predict a qualitatively correct trend.

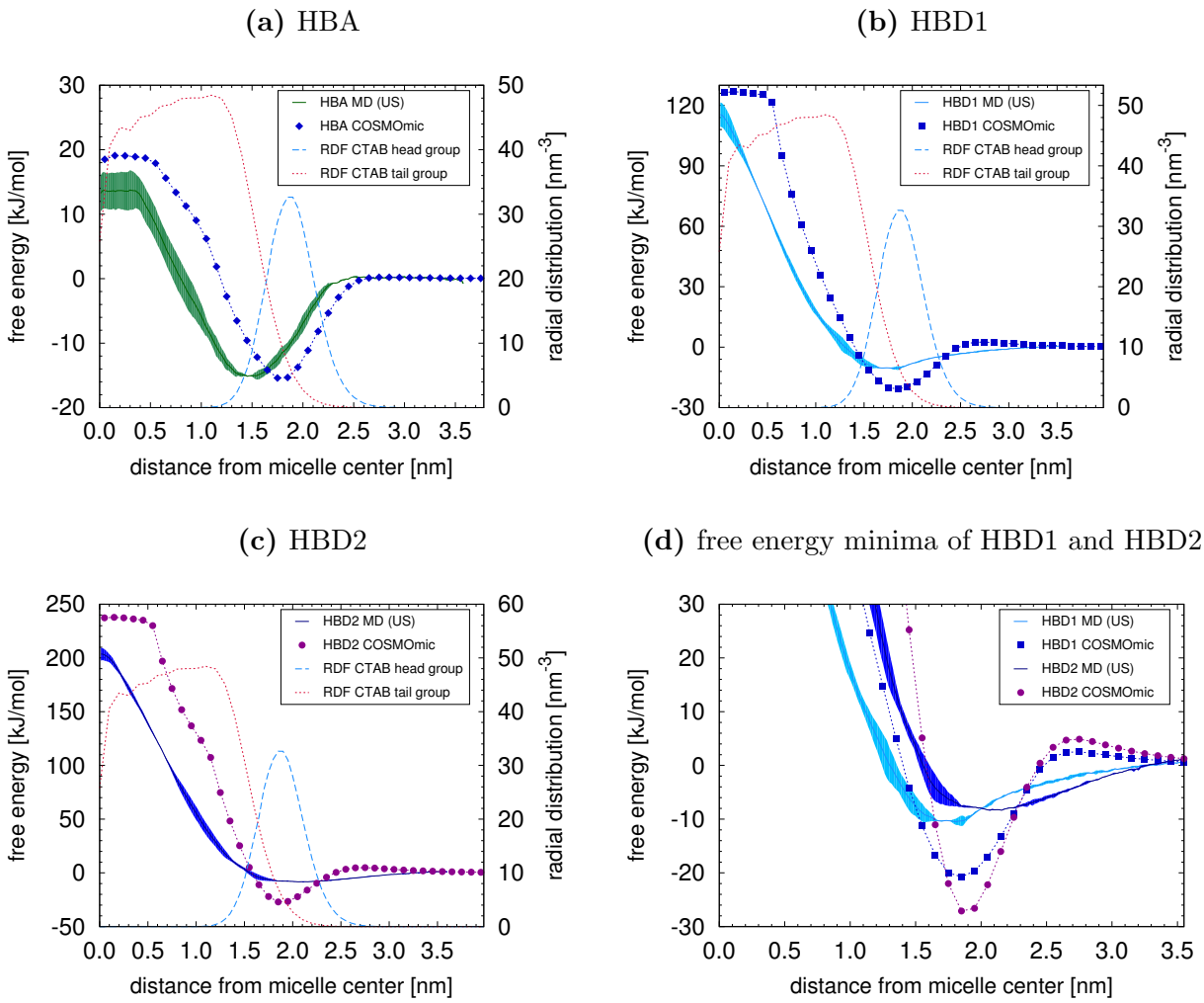

Figure 6.6: Free energy profiles of 4-hydroxybenzoic acid in neutral (HBA) and dissociated (HBD1 and HBD2) states in a CTAB micelle calculated from MD simulations (US method) and with COSMOmic. The colored areas show the error margin of the MD simulation results. Radial distribution functions of the surfactant head and tail group atoms calculated from MD simulation without solutes are plotted on the secondary y-axes.

Influence of CTAB micelle size on partition coefficient predictions with COSMOmic

The influence of micelle sizes (in terms of aggregetion numbers, N_{agg}) on partition coefficient predictions with COSMOmic in various micelles was analyzed in reference 206. It could be demonstrated that the calculations in CTAB are more sensitive on the micelle size in comparison to predictions in SDS and non-ionic micelles. Still, the usage of small micelles (aggregation number up to 50) was recommended, as more stable and reproducible results could be obtained compared to those calculated with larger micelles. However, only neutral solutes were considered. In this chapter, the influence of CTAB micelle size on the partition behavior predictions with COSMOmic is investigated with a focus on

charged solutes. Partition coefficients of 4-hydroxybenzoic acid, phloretic acid and 4-methoxyphenol in their neutral and charged states calculated with different CTAB micelle sizes are given in Table 6.5. As only influences on COSMOmic calculations are discussed, all partition coefficients were calculated with equation 2.74, as it is the equation used in the COSMOtherm implementation of COSMOmic. It can be seen that the calculated partition coefficients of charged molecules are highly sensitive to the micelle size. In general, with increasing aggregation number the partition coefficients decrease significantly. In the case of HBD1, PLD1 and MTD, the partition coefficients predicted with $N_{\text{agg}} \leq 40$ are relatively constant, but decrease for larger micelles. This behavior could be directly related to the changes in the free energy profiles. As an example, free energy profiles of HBD1 calculated using micelle sizes between 20 and 120 (step 10) are shown in Figure 6.7.

Table 6.5: Partition coefficients $\log K_i$ calculated with COSMOmic using different aggregation numbers of CTAB, all K_i have units of L/kg.

	Solute							
N_{agg}	HBA	HBD1	HBD2	PLA	PLD1	PLD2	MTP	MTD
20	2.36	3.52	4.77	2.66	3.89	5.28	2.68	4.11
30	2.50	3.52	4.67	2.79	3.91	5.21	2.76	4.14
40	2.60	3.46	4.45	2.90	3.85	5.02	2.85	4.11
50	2.62	3.26	4.06	2.91	3.63	4.63	2.84	3.92
60	2.61	2.97	3.50	2.88	3.30	4.05	2.80	3.63
70	2.58	2.71	3.02	2.83	3.01	3.56	2.75	3.36
80	2.54	2.51	2.41	2.76	2.64	2.92	2.69	3.02
90	2.49	2.14	1.97	2.69	2.37	2.44	2.59	2.70
100	2.40	1.81	1.45	2.57	2.02	1.85	2.47	2.28
110	2.32	1.45	0.87	2.49	1.66	1.23	2.37	1.83
120	2.16	0.85	0.01	2.32	1.08	0.33	2.22	1.07
exp.	2.64	3.91	-	2.94	3.99	-	2.38; 2.20	3.03

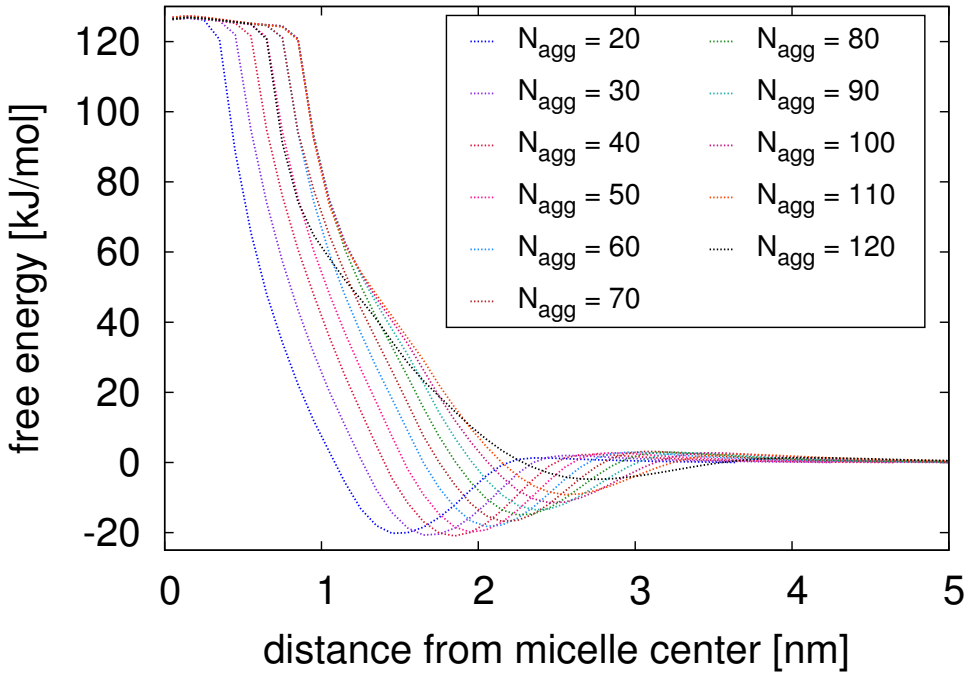

Figure 6.7: Free energy profiles of charged 4-hydroxybenzoic acid (HBD1) calculated with COSMOmic using different aggregation numbers of CTAB.

When using $N_{\text{agg}} \leq 40$ CTAB monomers, the depth of the minimum is relatively constant (changes are less than 0.5 kJ/mol). Differences are obtained only in the position of the minimum, which is related to the different system sizes. When using $N_{\text{agg}} \geq 50$ the depth of the minimum changes as well. This trend is even more distinct for the -2 charged molecules, which directly reflects in the calculated partition coefficients (higher differences are obtained). As expected, in the case of neutral solutes the trend is less pronounced. Thus, the stronger the electrostatic interactions between a micelle and a solute are, the greater the influence of the micelle size. Two main factors for this observation are: (i) the composition of the micelle at the free energy minimum and (ii) the shape of the micelle. The free energy minimum of all solutes (neutral and charged) is located in the hydrophilic head group region of CTAB. However, depending on the micelle size the head groups are differently mixed with the hydrophobic tails, that is the content of hydrophobic groups within the hydrophilic region changes. This can be seen from the radial distribution functions of the hydrophilic and hydrophobic part of micelles with different N_{agg} in Figure 6.8. With increasing aggregation number the overlapping area between hydrophobic and hydrophilic groups increases, that is more hydrophobic groups penetrate in the hydrophilic region. Therefore, with increasing micelle size weaker attractions between the CTAB head group and a charged solute is predicted, which leads to a higher free energy at the minimum and a lower partition coefficient.

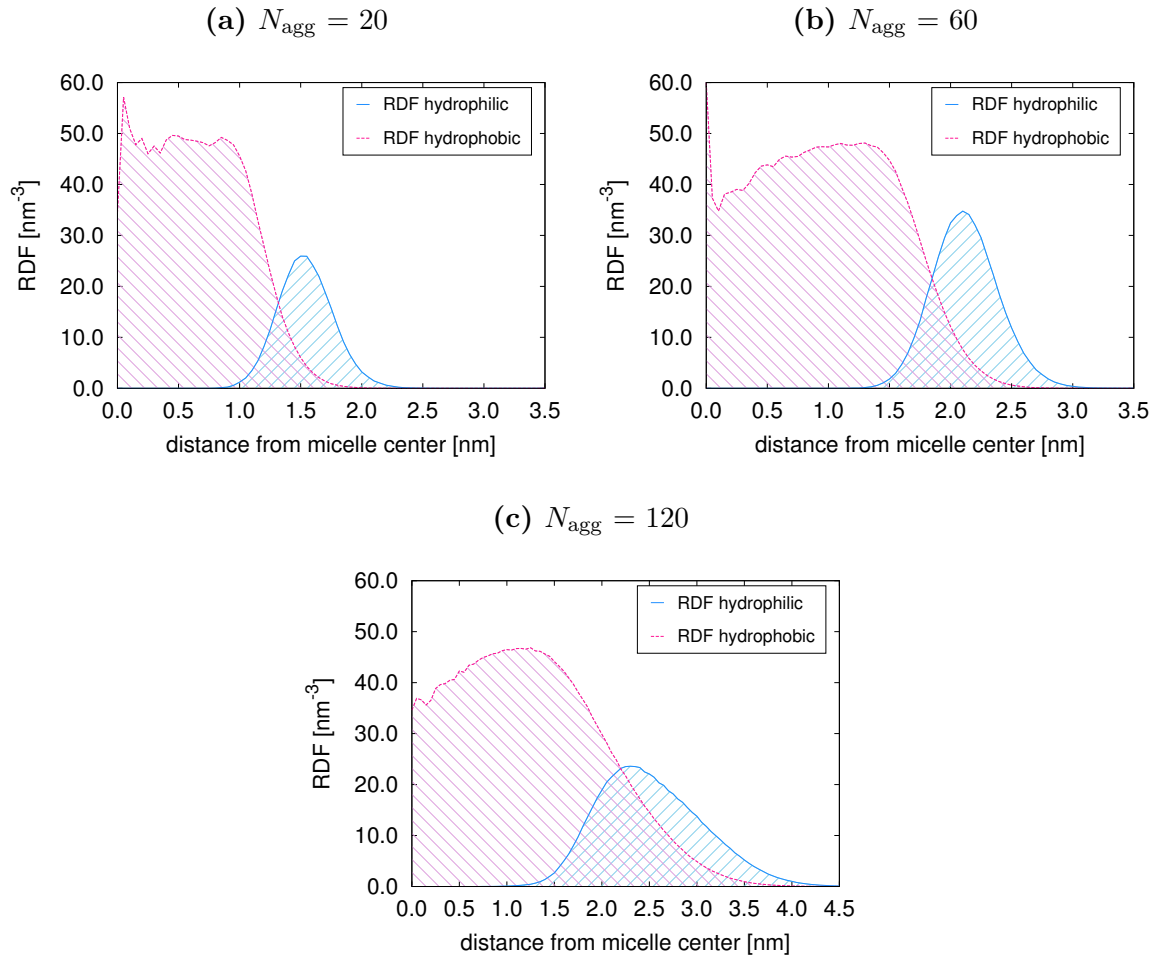

Figure 6.8: Radial distribution functions of the hydrophilic and hydrophobic part of CTAB micelles with different sizes.

The influence of the micelle shape (in terms of eccentricity, ε) for charged solutes is investigated. In Figure 6.9 the relation between the depth of the free energy minimum and the eccentricity is shown for different aggregation numbers. At first, the relation between eccentricity and aggregation number has to be evaluated. The most spherical micelles are those with aggregation numbers 40, 50 and 60 ($\varepsilon = 0.12$). Micelles with aggregation numbers 20 and 30 have slightly higher eccentricities (0.15 and 0.13, respectively). The same trend was observed in reference 206 where also the influence of higher concentrations and self-assembled micelles were studied. Thus, it could be assumed that aggregation numbers between 40 and 60 provide the most favorable geometrical arrangements for spherical CTAB aggregates. For micelles with $N_{agg} \geq 70$ an increase in the eccentricity is observed leading to the maximal value of $\varepsilon = 0.36$ for $N_{agg} = 120$.

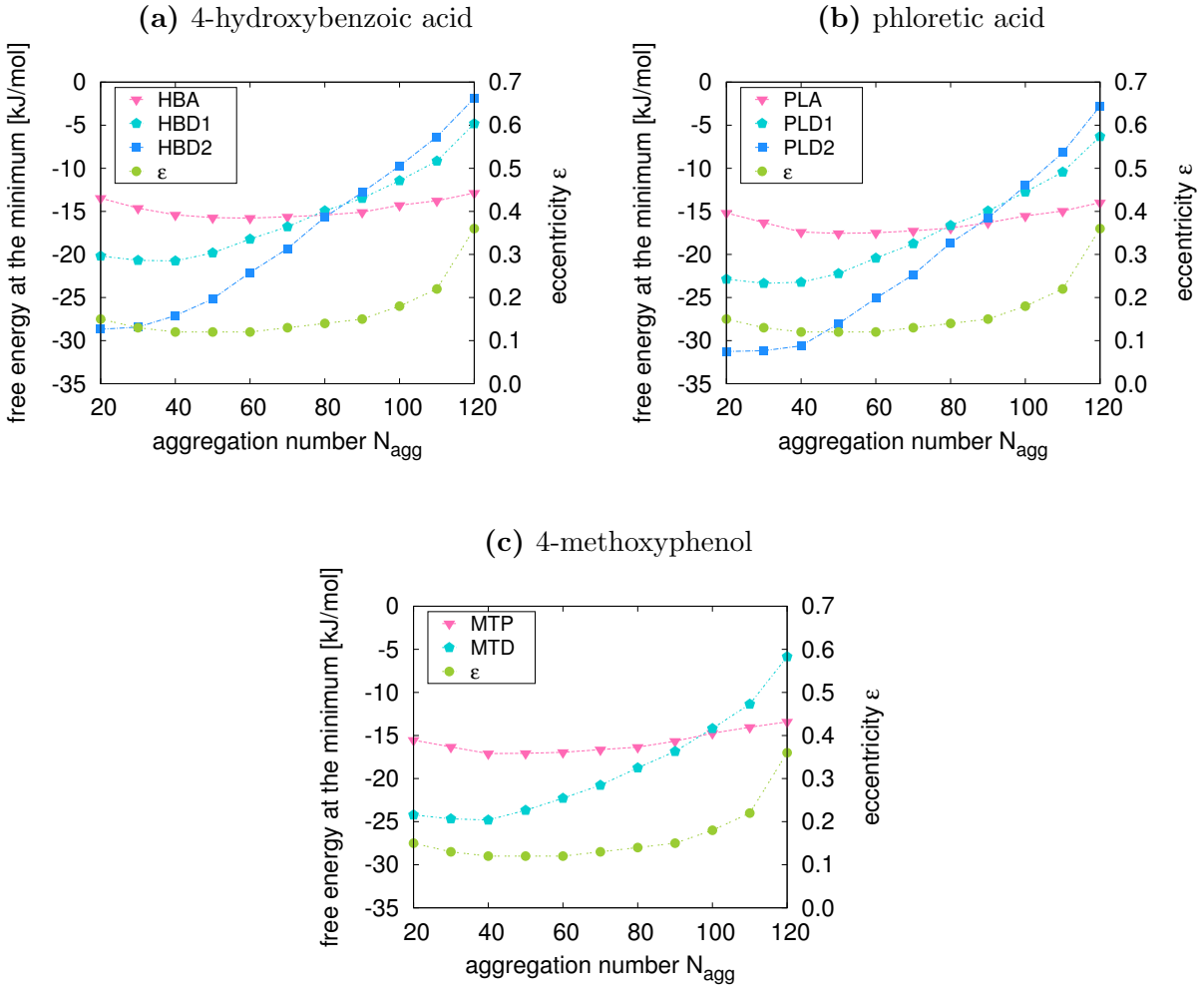

Figure 6.9: Free energy at the minimum in function of CTAB aggregation number. The eccentricity of the CTAB micelle is plotted on the secondary y-axes.

Interestingly, the same trend of the free energy at the minimum with N_{agg} is observed for all solutes and depends only on the charge of the solute (see Figure 6.9). The free energy at the minimum of solutes with charge -1 and -2 increases with N_{agg}. This behavior can be directly related to the change of the micelle composition in the head groups region (see Figure 6.8). Very small deviations (less than 0.5 kJ/mol) are obtained for $N_{\text{agg}} = 20 - 40$, which can be considered as negligible. Hence, for charged molecules the micelle composition is the key factor for the free energy at the minimum and therefore for the partition coefficient. Please notice that for $N_{\text{agg}} \geq 40$ both the eccentricity and the amount of hydrophobic groups in the head group region are increasing with N_{agg}.

It can be noticed from Figure 6.9 that the changes in the free energy at the minimum for neutral solutes are much less pronounced (maximal difference is ∼4 kJ/mol). The minimal free energies are obtained for the most spherical micelles (N_{agg} = 40-60). Higher free energies are calculated when using micelles with aggregation numbers 20 and 30, which can be attributed to their slighlty higher eccentricities. Further increase in the free energy is obtained for $N_{agg} \geq 70$ with maximal values calculated for $N_{agg} = 120$. Therefore, the usage of micelles with high aggregation numbers (higher eccentricity) for the COSMOmic calculation leads to lower partition coefficients. Although this trend is less pronounced for neutral compared to charged solutes, it still can have significant impact on the calculated partition coefficients (differences up to 0.6 log units).

It could be demonstrated that micelles with N_{agg} = 40-60 provide stable and reproducible results for both neutral and charged molecules. Hence, the findings support that small and spherical micelles are the most suitable for the calculations with COSMOmic. Please note that the focus of this analysis is not to determine which micelles show the best prediction quality compared to experimental results, but either which of them provide reliable and to the most possible extend unbiased and reproducible results. Actually, each aggregation number could be "accidentally" the best for a particular micelle-solute combination. In the future, it would be interesting to determine the influence of the aggregation number in MD simulations of solute partitioning. It can be assumed that it is much less pronounced compared to COSMOmic calculations, as COSMOmic relies on spherical micelles. In COSMOmic the structure of the micelles is obtained in a simulation without solute and cannot change according to the presence of the solute, whereas in MD simulations the micelle structure is dynamic.

Lipophilicity profiles

In order to compare the prediction quality of COSMOmic for the partitioning of molecules in their neutral and charged state, lipophilicity profiles of six solutes in CTAB micelles are calculated and compared to experimental data. The experimental partition coefficients are taken from reference 105 and are determined with the MLC technique. The lipophilicity profiles are shown in Figure 6.10 and the partition coefficients of the solutes in their neutral and dissociated state are given in Table 6.6.

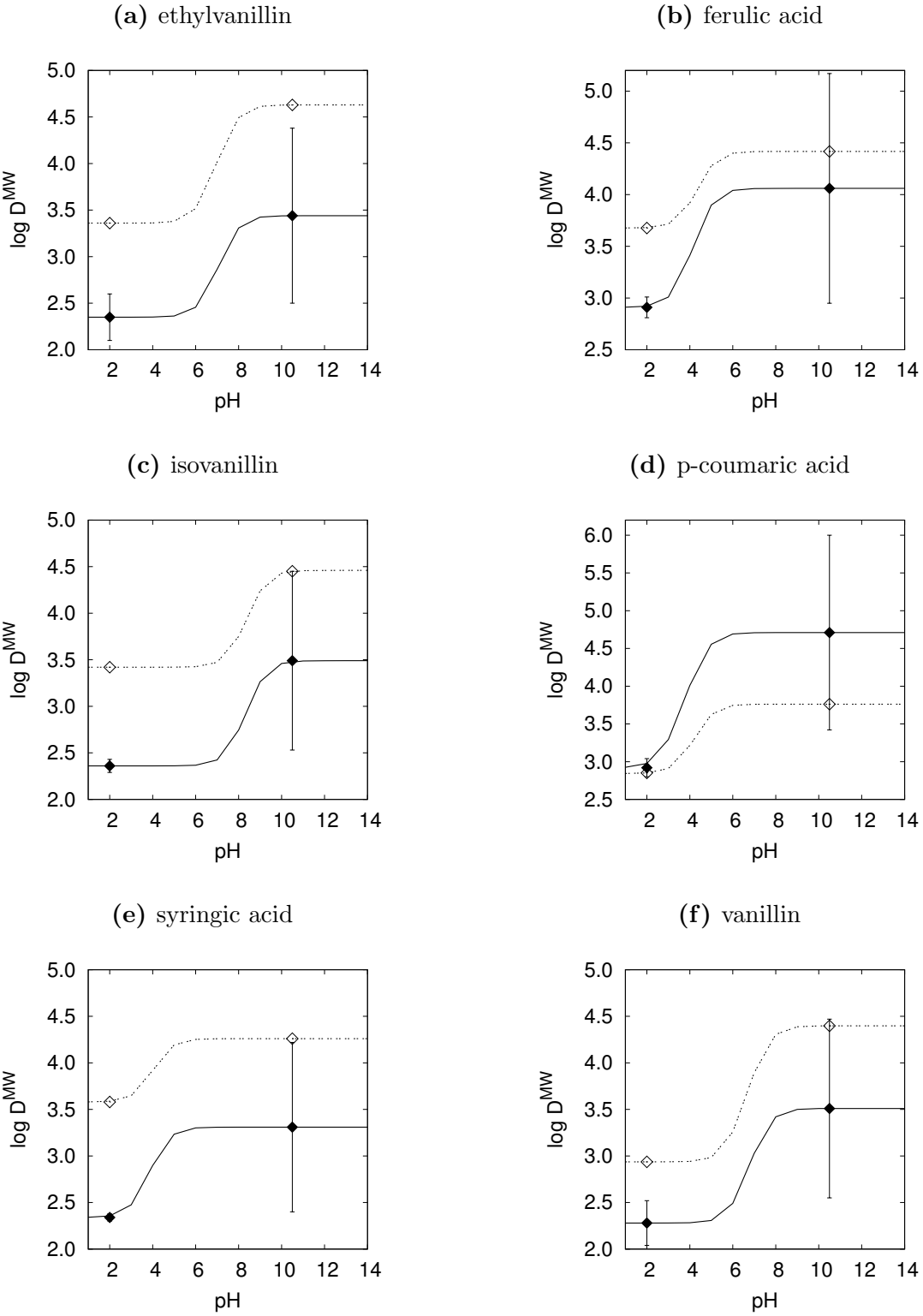

Figure 6.10: Lipophilicity profiles of six solutes in CTAB micelles, where the solid lines are experimental data and the dashed lines are COSMOmic calculations. The experimental data is taken from reference 105, where the determination with MLC is performed at pH = 2 of the mobile phase for the neutral state of the molecules and at pH = 10.5 for the dissociated state, respectively.

Table 6.6: Partition coefficients of neutral and ionized solutes in CTAB micelles calculated with COSMOmic in comparison with experimental data.[105] All K_i have units of L/kg.

Solute	neutral		ionized	
	$\log K_i^{\text{calc}}$	$\log K_i^{\text{exp}}$	$\log K_i^{\text{calc}}$	$\log K_i^{\text{exp}}$
Ethylvanillin	3.36	2.35 ± 0.25	4.63	3.44 ± 0.94
Ferulic acid	3.68	2.91 ± 0.10	4.42	4.06 ± 1.11
Isovanillin	3.42	2.36 ± 0.07	4.45	3.49 ± 0.96
p-Coumaric acid	2.85	2.92 ± 0.12	3.76	4.71 ± 1.29
Syringic acid	3.58	2.34 ± 0.04	4.26	3.31 ± 0.91
Vanillin	2.94	2.28 ± 0.24	4.40	3.51 ± 0.96
RMSE	0.89		0.92	

It can be seen that the prediction quality of COSMOmic for neutral and dissociated solutes in CTAB micelles is similar with an overall RMSE of 0.89 for neutral molecules and 0.92 for negatively charged, respectively. The absolute deviation between calculated and experimentally determined partition coefficients for the neutral and the dissociated state of each molecule is comparable for all solutes except p-coumaric acid, where a very good agreement is observed for the neutral state (an absolute deviation of 0.07 log units) and a much higher deviation for the negatively charged state (an absolute deviation of 0.95 log units). p-Coumaric acid is also the only solute, for which the predicted partition coefficients underestimate the experimentally determined, whereas for the other five solutes the predictions overestimate the experimental values. As already discussed in chapter 6.1 the partition coefficient determination of charged solutes with MLC is challenging, which leads to higher error bars in the case of charged molecules. The predicted partition coefficients of the dissociated state are within the range of the experimental error for five of the six considered solutes (except ethylvanillin). However, the accuracy of the COSMOmic calculations in CTAB micelles is significantly lower compared to those in nonionic Triton X micelles (see chapter 5.2.3) and in anionic SDS micelles (see chapter 6.3.3). It was observed that the predictions for CTAB micelles are more sensitive to the micelle structures and show generally higher deviations from experimental data than nonionic surfactants and SDS.[206] Since the deviation for the neutral and charged state is similar, the lower prediction quality in CTAB can be attributed to the interactions between CTAB and the solutes in general. It can be concluded that the COSMOmic prediction quality for neutral and charged solutes in CTAB micelles is comparable and no decline due to the charge of the solutes is observed. However, only six solutes are considered in this comparison. For a final conclusion, a larger set of partition coefficients in

CTAB is necessary.

6.4 Summary

In this chapter the solute partitioning was investigated in four types of micelles. Free energy profiles calculated from MD simulations with the US method and with COSMOmic were compared. For a quantitative validation, partition coefficients predicted with both methods were compared to experimental results. The partition behavior of negatively charged solutes in zwitterionic and cationic micelles was considered. Due to lack of experimental data for charged solutes in the literature, partition coefficients of dissociated solutes in CTAB micelles were measured using micellar liquid chromatography. It could be demonstrated that that the determination with MLC can be challenging for charged solutes, which corresponds to previous observations.[42,105,221]

The main findings of the investigation on partition equilibria predictions using MD simulations with the US method and the COSMOmic model can be summarized as follows. As micelles have a higher water content in comparison to membranes (see also chapter 6.3), it could be demonstrated that the equation used for the calculation of partition coefficients from free energy profiles has a higher impact for micelles than for membranes.[161] On average equation 2.74 (with water correction) results in $\log K_i$ which are 0.5 higher than calculations with equation 2.61 (using a defined border between the micellar and the aqueous phase). Comparison to experimental values gives no clear trend which equation leads to better predictions (see Table 6.4). In any case, if two models are compared, the same equation should be used for both methods. If compared to experimental data, the choice of the equation should depend on the experimental technique and one should ask the question how the water in the micelle phase was treated during the experiment and the analysis.

In the nonionic TX114 micelle the partition behavior of neutral 4-hydroxybenzoic acid (HBA) was investigated. For this system both MD and COSMOmic provide similar free energy profiles according to the position and depth of the minimum, which show that the solute is favorably located in the hydrophilic region of TX114. The predicted partition coefficients are in good agreement with experimental results. The partition behavior of 3-phenylpropanoic acid in the neutral (PPA) and dissociated (PPD) form was calculated in the zwitterionic HePC micelle. Some deviations in the free energy profiles from MD simulations and COSMOmic are obtained in the hydrophobic core and in the transition region. By evaluating the free energy in the hydrophobic core by cyclohexane/water[235] transfer free energies, the US

profiles could be identified to be more reliable in this region. Due to lack of experimental data the calculated partition coefficients could be evaluated only qualitatively, both methods predict the expected trend (decreasing partition coefficient when the solute is charged). The partitioning of 4-methoxyphenol (MTP) and naphthalene (NAPH) in the anionic SDS micelle was investigated. Analog to the zwitterionic micelle, the free energy profiles from MD simulations and COSMOmic show differences in the hydrophobic core and in the transition region. Again, by comparison to cyclohexane/water transfer free energies the US profiles could be considered as more reliable. The predicted partition coefficients of NAPH are in good agreement with experimental data, whereas both models overestimate the partition coefficient of MTP. 4-Hydroxybenzoic acid in the neutral (HBA) and charged (HBD1 and HBD2) state were investigated in cationic CTAB micelles. Similar free energy profiles for the neutral form are obtained, the predicted partition coefficients are in good agreement with experimental data. In the case of the dissociated states, only the COSMOmic model predicts the expected trend (increasing partition coefficient with increasing negative charge of the solute). In contrast, the US method is not able to predict an increase in the partition coefficient when the solutes are negatively charged, which probably can be attributed to the electrostatic model in the force field used for the MD simulations. For this system, the usage of polarizable force fields could be advantageous. In general, the results from MD simulations and COSMOmic are comparable. Additionaly, lipophilicity profiles of six ionizable solutes in CTAB micelles were calculated with COSMOmic and compared with experimental results determined with MLC. The COSMOmic model is able to predict partition coefficients of neutral and ionized solutes in CTAB micelles with comparable accuracy, but deviations to experimental results are observed for both solutes.

In the next chapter MD simulations and the COSMOmic model will be further evaluated for partition equilibria predictions in mixed micellar systems.

7 Micellization and partition equilibria in mixed nonionic/ionic micellar systems

In the previous chapter, the partition behavior of various solutes in micelles of all surfactant types (nonionic, zwitterionic, anionic and cationic) was investigated with MD simulations and COSMOmic. However, only micelles consisting of one surfactant type were considered. For practical applications, mixed surfactant solutions are also relevant. In this chapter, the micellization process of mixed nonionic/ionic surfactant systems is investigated with MD self-assembly simulations. Furthermore, COSMO-RS and COSMOmic are evaluated and compared for the prediction of partition equilibria in various mixed nonionic/ionic micelles. The results of the following sections were in parts published as Yordanova, D.; Ritter, E.; Smirnova, I.; Jakobtorweihen, S. **Micellization and partition equilibria in mixed nonionic/ionic micellar systems: Predictions with molecular models** in Langmuir.[296]

7.1 TX114/SDS and TX114/CTAB mixed micelles

In this chapter, self-assembly MD simulations of mixed nonionic/ionic surfactant systems were performed to investigate the aggregation process at different surfactant ratios, as well as the composition of the mixed micelles. Moreover, micelle structures taken from these simulations are used for calculation of partition equilibria in chapter 7.1.3. Please note that preassembled micelles can also be used. In reference 206, the prediction quality of COSMOmic using self-assembled and preassembled micelles was compared. It could be demonstrated that self-assembled and preassembled micelles with similar aggregation numbers and atomic distributions lead to similar predictions with COSMOmic. Therefore, the focus of the self-assembly simulations is not to obtain micelle structures, but to investigate the aggregation behavior in mixed surfactant systems.

7.1.1 Simulations details

Initial configurations for the self-assembly simulations of TX114/SDS and TX114/CTAB were prepared with PACKMOL[275] containing 200 randomly placed surfactant monomers. A concentration of 0.13 M was adjusted by adding a specific amount of water molecules. Counter ions were added explicitly to ob-

tain charge neutrality. An energy minimization was performed with the steepest descent method. First, a 600 ps equilibration step was carried out in the NVT ensemble with a time step of 1 fs. Afterwards, a simulation in the NPT ensemble with the Berendsen barostat[247] was performed for 200 ps. Finally, the systems were simulated for 200 ns in the NPT ensemble with the Parinello-Rahman barostat[248] (isotropic coupling constant $\tau_p = 2$ ps) with a time step of 2 fs. The temperature was fixed in all simulations at 283 K using the Nosé-Hoover thermostat[246] (coupling constant $\tau_t = 1$ ps).

7.1.2 Formation of mixed micelles: Self-assembly MD simulations

Self-assembly simulations of TX114/SDS mixtures at three molar ratios ($\alpha = 0.1$, $\alpha = 0.5$ and $\alpha = 0.9$) and TX114/CTAB ($\alpha = 0.5$) are performed. In Figure 7.1 the progress of the maximum aggregation numbers over the simulation time is shown. It can be seen that micelles become more stable when the content of the ionic surfactant increases, whereas more fluctuations in the aggregation number are obtained in the system with higher TX114 content. Moreover, the aggregation numbers decrease with increasing content of ionic surfactant. This behavior could be explained with the micelle size distributions in micellar solutions. It is known that all micellar systems exhibit some degree of polydispersity with respect to micelle aggregation numbers.[297] This is characteristic especially for solutions of nonionic surfactants.[175,262,263] However, in case of mixed nonionic/ionic micelles monodisperse solutions are observed. Kamayama and Takagi[298] reported monodisperse $C_{12}E_8$/SDS mixed micelles obtained from electrophoretic light scattering. Furthermore, Tokiwa and Aigami[299] investigated the micelle sizes of mixed SDS with dodecyl polyoxyethylene ether and determined decreasing micelle sizes with increasing SDS content. The same trend was observed for TX100/SDS mixtures at low ionic strength,[300] which is in agreement with the trend from the self-assembly simulations in this work. In contrast, Komaromy-Hiller et al.[301] reported no significant influence on the aggregation number of TX114 when SDS and CTAB surfactants are added. Nevertheless, this can be attributed to the concentrations of the ionic surfactants in their study, which are very low (below their CMCs).

Furthermore, at different nonionic/ionic surfactant ratios the obtained aggregation numbers differ significantly, as expected. The aggregation number of TX114/SDS micelles is 24 for $\alpha_{\mathrm{SDS}} = 0.9$ and 200 for $\alpha_{\mathrm{SDS}} = 0.1$, respectively. Both are in the range of aggregation numbers, obtained from self-assembly MD simulations in the literature: 20-38[185] for pure SDS and 183-216[173] for pure TX114. The aggregation numbers in the equimolar TX114/SDS and TX114/CTAB systems are 50 and 79, respectively. Both are higher than those

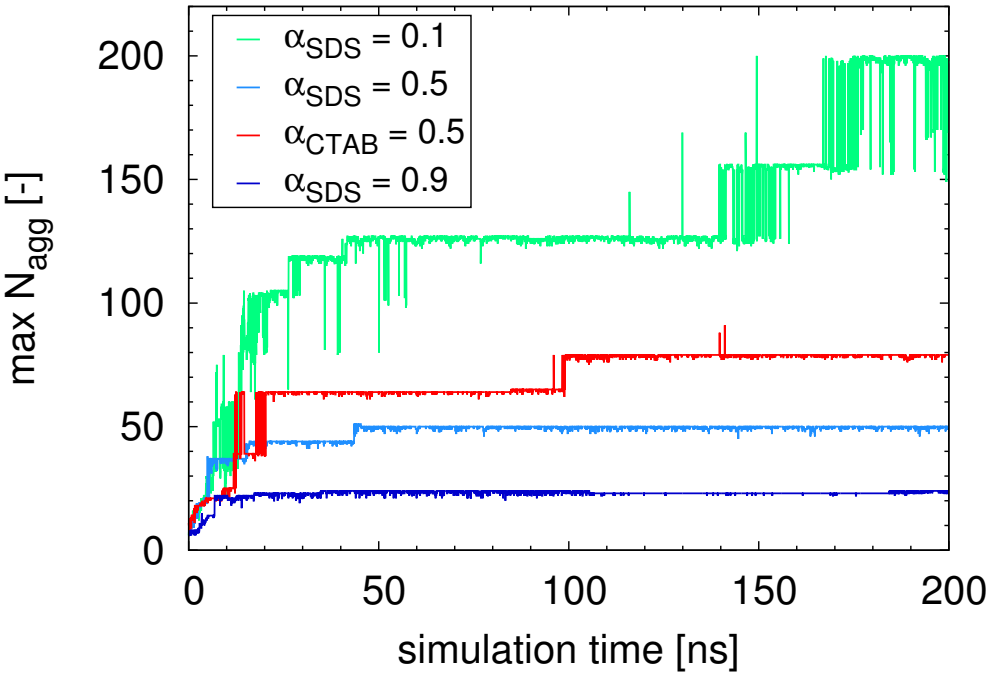

Figure 7.1: Maximum aggregation number over time from self-assembly simulations of TX114/SDS and TX114/CTAB systems at different compositions α.

of pure SDS and CTAB micelles (20-38[185] and 23[185], respectively), which can be attributed to the TX114 content. Therefore, the expected trend could be observed.

It has been demonstrated that the composition in the mixed micelles is a crucial parameter for the prediction of partition equilibria in mixed micelles.[42] In all self-assembly simulations in this work, the composition in the mixed micelles approaches the composition in the system, where the compositions in the micelles are calculated from the micelles formed at the end of the 200 ns simulations (data not shown). This is in agreement with the experimental observations for TX100/CTAB mixtures at concentrations higher than ∼2 mM, where the composition of the mixed micelles is equal to the composition in the solution.[213] The same behavior was observed for Brij35/CTAB, where the composition of the mixed micelles is equal to the composition in the solution at concentrations higher than ∼1 mM.[94] Please note that this is valid at concentrations higher than the CMCs of the surfactants. At concentrations below their CMCs, the composition of the mixed micelles differs from the surfactant ratios in the solution.[94,213] As the relevant concentrations in both simulations and experiments are higher than the CMCs of the mixed micellar solutions, for the prediction of partition equilibria in this work the composition in the mixed micelles is assumed to be equal to the composition in the initial mixed surfactant solution.

7.1.3 Partition equilibria: Comparison between experiments and predictions

Partition coefficients of neutral solutes in mixed TX114/SDS and TX114/CTAB micelles are experimentally determined using MLC and predicted with the COSMO-RS model and COSMOmic. For both systems equimolar solutions are used in the experiments and therefore also in the calculations. In this chapter, all partition coefficients were measured only once. However, the MLC technique has been established for the determination of partition coefficients of neutral molecules and the same method and retention model as described in the literature[42,105] are used. Therefore, the maximum error is assumed to be $\Delta \log K_i = \pm 0.25$ in accordance with previous experiments. The experimentally obtained and calculated partition coefficients in the system TX114/SDS are shown in Table 7.1.

Table 7.1: Calculated partition coefficients of neutral solutes in TX114/SDS mixed micelles (equimolar composition) in comparison with experimental data. All K_i have units of L/kg.

Solute	$\log K_i^{\text{calc}}$		$\log K_i^{\text{exp}}$
	COSMOmic[a]	COSMO-RS	
4-Hydroxybenzaldehyd	1.63	1.27	1.84
4-Hydroxybenzoic acid	2.24	2.46	1.97
Coumarin	1.56	0.88	1.68
Ferulic acid	2.63	2.84	2.25
Isovanillin	1.70	1.26	1.76
p-Coumaric acid	2.30	2.60	2.25
Syringic acid	2.34	2.60	1.97
Vanillic acid	2.42	2.64	2.09
Vanillin	1.86	1.51	1.72
RMSE	0.25	0.55	

[a]The micelle structure used for the COSMOmic calculation was taken from a 200 ns self-assembly simulation, $N_{\text{agg}} = 50$.

It can be seen that the predictions with COSMOmic are in very good agreement with the experimental data (RMSE = 0.25). The COSMO-RS model shows higher deviations with an overall RMSE = 0.55. The same relation is obtained for the system TX114/CTAB, where the RMSE for the COSMOmic calculation is 0.37 and 0.64 for COSMO-RS, respectively (see Table 7.2). Where for the system TX114/SDS no trend is observed, all COSMOmic calculations for TX114/CTAB with one exception (4-hydroxybenzoic acid), show an overprediction. If this is a general behavior for the TX114/CTAB system, cannot be elucidated here as the number of solutes is too small. However, when

comparing the two systems TX114/SDS and TX114/CTAB, it can be noticed that both COSMOmic and COSMO-RS have slightly higher deviations in the TX114/CTAB micelles compared to TX114/SDS. Since with a few exceptions the same set of solutes is considered, the higher deviation can be attributed to the CTAB surfactant. Higher deviations and an overprediction were also observed in pure CTAB micelles (see chapter 6.3.4).

Table 7.2: Calculated partition coefficients of neutral solutes in TX114/CTAB mixed micelles (equimolar composition) in comparison with experimental data. All K_i have units of L/kg.

Solute	$\log K_i^{\text{calc}}$		$\log K_i^{\text{exp}}$
	COSMOmic[a]	COSMO-RS	
4-Hydroxybenzaldehyd	1.61	2.87	1.82
4-Hydroxybenzoic acid	2.10	1.55	1.93
Acetophenone	2.00	1.14	1.54
Coumarin	1.89	1.06	1.63
Ethylvanillin	2.37	2.03	1.92
Ferulic acid	2.80	3.28	2.38
Isovanillin	2.00	1.53	1.63
Orthovanillin	2.23	1.62	1.67
Phenol	2.14	2.26	1.81
Resorcin	1.95	2.78	1.86
Vanillic acid	2.47	3.05	1.92
Vanillin	2.07	1.77	1.84
RMSE	0.37	0.64	

[a]The micelle structure used for the COSMOmic calculation was taken from a 200 ns self-assembly simulation, $N_{\text{agg}} = 64$.

It can be concluded that the COSMO-RS pseudo phase approach partition coefficient calculations are in reasonable agreement to the experimental data. Nevertheless, the usage of the COSMOmic model is beneficial, as the overall RMSEs in both systems are ∼0.3 log units lower than with COSMO-RS (pseudo phase). Hence, taking into account the anisotropy of the micelles, which is neglected in the pseudo phase approach, is of importance for the prediction quality of partition equilibria in mixed micellar systems.

7.2 Brij35/CTAB mixed micelles

For mixed micellar systems not only equimolar mixture are important, rather different compositions are relevant for the design of new processes. In order to evaluate COSMO-RS and COSMOmic for the prediction of partition equilibria in mixed micelles, the consideration of micelles with different compositions

is of importance. In this chapter, partition coefficients of neutral solutes in Brij35/CTAB mixed micelles at different molar ratio of CTAB are calculated, as experimental data is available in reference 42.

7.2.1 Simulation details

For the Brij35/CTAB mixed micelles, preassembled micelles were used as starting structures. They were generated with PACKMOL,[275] whereas water molecules were added such that the system had a surfactant concentration of 0.08 M. $N_{\mathrm{agg}} = 40$ was chosen for all micelles. After an energy minimization with the steepest descent method, two equilibration steps were performed: a 600 ps simulation in the NVT ensemble, followed by a 200 ps simulation in the NPT ensemble with the Berendsen barostat[247]. The NPT simulations for sampling were carried out for 40 ns with the Parinello-Rahman barostat[248] (isotropic coupling constant $\tau_p = 2$ ps) and the Nosé-Hoover thermostat[246] ($\tau_t = 1$ ps). The temperature was fixed at 298 K and a time step of 2 fs was used.

7.2.2 Partition coefficients of neutral solutes

The predicted partition coefficients in comparison to experimental data are shown in Figure 7.2 and listed in Table 7.3. It is known that usually higher partition coefficients are obtained for CTAB than for nonionic surfactants due to enhanced electrostatic interactions between the solutes and the cationic head group of CTAB. However, the partition coefficients do not increase linear with increasing CTAB content, which can be seen from the experimental data in Figure 7.2. In general, it can be observed that the trend predicted by COSMOmic is qualitatively in agreement with the experiments. For some solutes (coumarin, ferulic acid, isovanillin) the predictions of COSMOmic partition coefficients are also quantitatively in good agreement to the experimental values at low CTAB content ($\alpha \leq 0.5$). Higher deviations are obtained at higher CTAB content, where the predicted partition coefficients overestimate the experimental values. An increase in the partition coefficients in CTAB compared to those in Brij35 is predicted by COSMO-RS as well. However, in the COSMO-RS calculations the partition coefficients increase linear with increasing CTAB content, which is not in accordance with the experimental data. The obtained linear trend can be attributed to the pseudo phase approach (micelle structure not taken into account) in COSMO-RS, whereas the trend predicted by COSMOmic is more precise.

It can be noticed that the deviations between the COSMO-RS calculations and experimentally determined partition coefficients increase with increasing CTAB content as well. The overall RMSE of COSMOmic and COSMO-RS in dependence of the CTAB content is shown in Figure 7.3. The RMSE of COSMO-RS increases from 0.56 in pure Brij35 micelle to 0.73 in pure CTAB, whereas the difference is more pronounced for COSMOmic (from 0.35 to 0.90, respectively). Still, for $\alpha \leq 0.8$ COSMOmic provides quantitatively better predictions than COSMO-RS. This is not the case in pure CTAB, where the overall RMSE of COSMO-RS is 0.17 log units lower than for COSMOmic. However, it has to be noted that the sensitivity of the COSMOmic calculations on the used micelle structure and size increases with increasing CTAB content.[206] In chapter 6.3.4 the influence of the CTAB micelle size on the COSMOmic calculations was investigated in detail. One finding was that the CTAB micelle size has a significant influence especially for charged solutes, but also for neutral molecules deviations up to 0.6 log units were obtained.

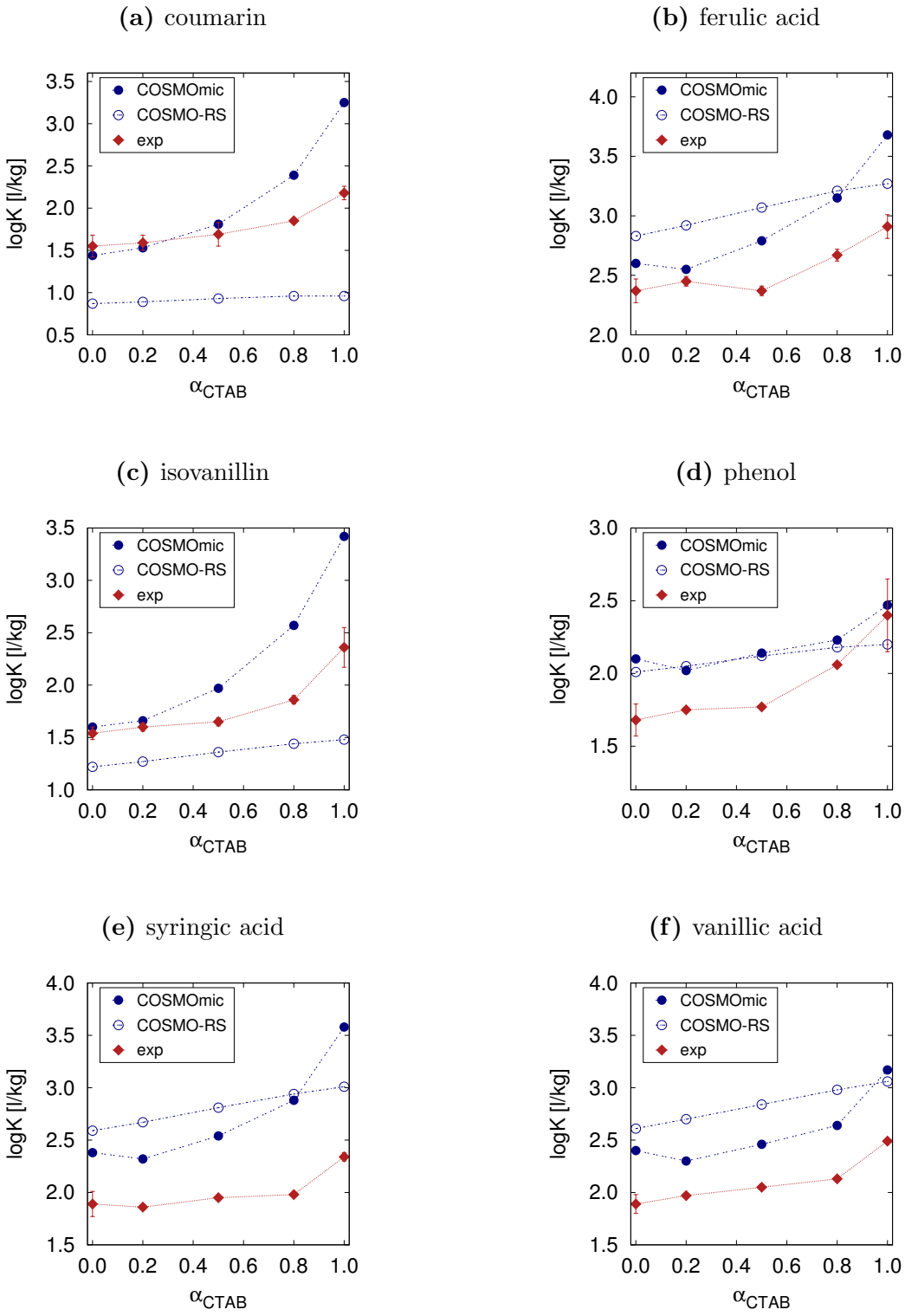

Figure 7.2: Partition coefficients of 6 neutral solutes in Brij35/CTAB mixed micelles at different micelle compositions. The experimental data is taken from reference 42. Lines are added as guides for the eye.

Table 7.3: Calculated and experimentally determined partition coefficients of neutral solutes in Brij35/CTAB mixed micelles at different micelle composition, where α is the mole fraction of CTAB, $\log K_i^{\mathrm{mic}}$ is the partition coefficient calculated with COSMOmic and $\log K_i^{\mathrm{RS}}$ with COSMO-RS, respectively. The experimental data is taken from reference 42. All partition coefficients have units of L/kg.

Solute	$\alpha = 0.0$			$\alpha = 0.2$			$\alpha = 0.5$			$\alpha = 0.8$			$\alpha = 1.0$		
	$\log K_i^{\mathrm{mic}}$	$\log K_i^{\mathrm{RS}}$	$\log K_i^{\mathrm{exp}}$	$\log K_i^{\mathrm{mic}}$	$\log K_i^{\mathrm{RS}}$	$\log K_i^{\mathrm{exp}}$	$\log K_i^{\mathrm{mic}}$	$\log K_i^{\mathrm{RS}}$	$\log K_i^{\mathrm{exp}}$	$\log K_i^{\mathrm{mic}}$	$\log K_i^{\mathrm{RS}}$	$\log K_i^{\mathrm{exp}}$	$\log K_i^{\mathrm{mic}}$	$\log K_i^{\mathrm{RS}}$	$\log K_i^{\mathrm{exp}}$
Coumarin	1.44	0.87	1.55 ± 0.13	1.53	0.89	1.59 ± 0.09	1.81	0.93	1.69 ± 0.14	2.39	0.96	1.85 ± 0.03	3.25	0.96	2.18 ± 0.08
Ferulic acid	2.60	2.83	2.37 ± 0.10	2.55	2.92	2.45 ± 0.04	2.79	3.07	2.37 ± 0.04	3.15	3.21	2.67 ± 0.05	3.68	3.27	2.91 ± 0.10
Isovanillin	1.60	1.22	1.54 ± 0.06	1.66	1.27	1.60 ± 0.04	1.97	1.36	1.65 ± 0.04	2.57	1.44	1.86 ± 0.04	3.42	1.48	2.36 ± 0.19
Phenol	2.10	2.01	1.68 ± 0.11	2.02	2.05	1.75 ± 0.01	2.14	2.12	1.77 ± 0.01	2.23	2.18	2.06 ± 0.01	2.47	2.20	2.40 ± 0.25
Syringic acid	2.38	2.59	1.89 ± 0.12	2.32	2.67	1.86 ± 0.02	2.54	2.81	1.95 ± 0.02	2.88	2.94	1.98 ± 0.02	3.58	3.01	2.34 ± 0.04
Vanillic acid	2.40	2.01	1.89 ± 0.09	2.30	2.70	1.97 ± 0.02	2.46	2.84	2.05 ± 0.02	2.64	2.98	2.13 ± 0.02	3.17	3.06	2.49 ± 0.02
RMSE	0.35	0.56		0.26	0.59		0.37	0.64		0.59	0.70		0.90	0.73	

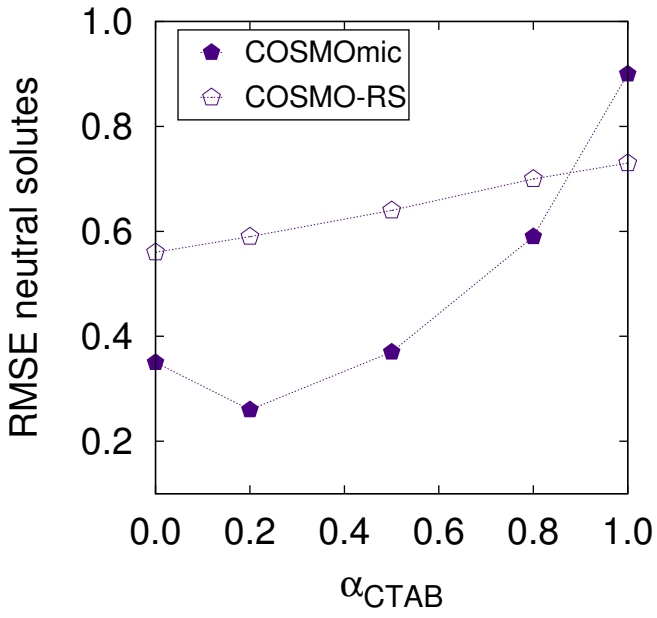

Figure 7.3: RMSE of COSMOmic and COSMO-RS calculated partition coefficients of 6 neutral solutes in Brij35/CTAB mixed micelles at different surfactant compositions. Lines are added as guides for the eye.

In summary, although the COSMO-RS predictions are reasonable, the predictions with COSMOmic, where the anisotropic environment within the micelle is accounted for, can be considered as more precise and are quantitatively in better agreement with experimental results. Hence, it is important to take into account the three-dimensional structure of the micelles.

7.2.3 Partition coefficients of ionized solute

It is known that the partition behavior in mixed nonionic/ionic micelles can be influenced by adjusting the pH of the surfactant solution. A partition coefficient increase in Brij35/CTAB mixed micelles is obtained when the acidic solute is dissociated. This effect is due to attractive forces between the negatively charged solute and the positively charged head group of CTAB. Therefore, it is expected that the partition coefficient increases with increasing CTAB content in the mixed micelle for acidic solutes. Experimentally determined partition coefficients of ionized isovanillin in Brij35/CTAB mixed micelles at different molar ratios of CTAB are available in the literature[42] and calculated in this chapter using the COSMOmic model. For these calculations the protolytic equilibria not only in water, but also in the anisotropic micelle based on free energy profiles calculated with COSMOmic are taken into account. This is possible by applying a thermodynamic cycle (see chapter 4.5). The pK_a value of 8.89 for isovanillin in water is used in the calculations.[302] As an example, free

energy profiles of isovanillin (neutral and charged) in a pure CTAB micelle and the estimated position dependent pK_a are shown in Figure 7.4.

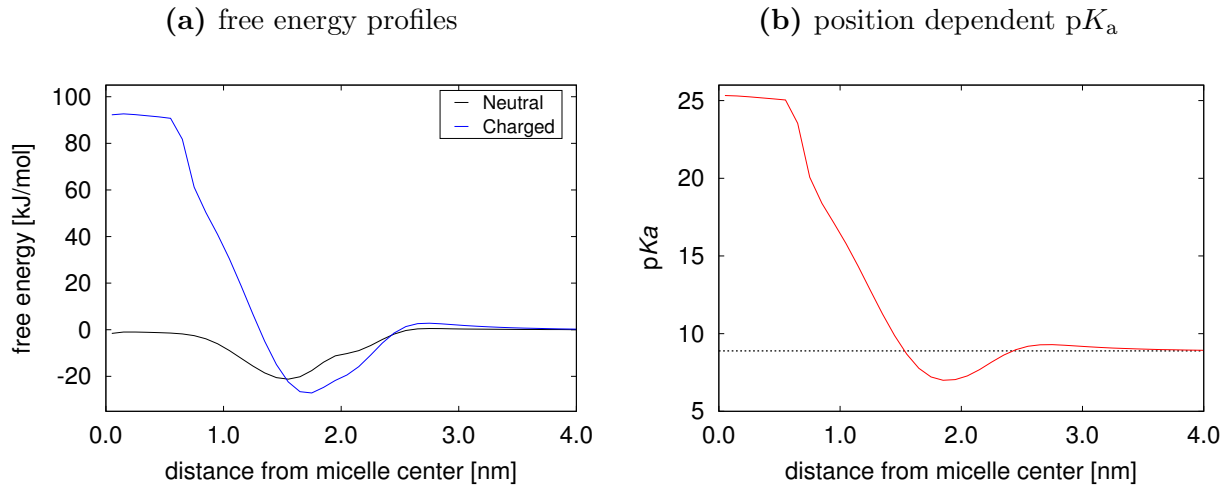

Figure 7.4: (a) Free energy profiles of neutral and charged isovanillin in a CTAB micelle calculated with COSMOmic. (b) Calculated position dependent pK_a of isovanillin in a CTAB micelle. The dotted line represents the experimental pK_a of isovanillin in water.[302]

In order to distinguish between the micellar and water phase, the phase boundary between them has to be defined. The free energy profiles of neutral isovanillin obtained with COSMOmic were used to define a border between both phases accordingly: the water phase begins where the free energy has reached a value near zero (maximal deviation < 0.40 kJ/mol). Thereby, the interface positions are obtained, which are 2.65 nm for pure CTAB micelles, 2.75 nm for $\alpha = 0.8$, 3.35 nm for $\alpha = 0.5$ and 4.05 nm for $\alpha = 0.2$, respectively.

The calculated and experimental partition coefficients of ionized isovanillin in Brij35/CTAB mixed micelles are given in Table 7.4. The experimental partition coefficients were determined with MLC at pH = 10.5.[42] As the determination of partition coefficients with MLC is based on retention models, it is not possible to distinguish between different ionization states of the same molecule and the ionization state of the solutes is calculated according to the adjusted pH of the mobile phase. Therefore, this procedure does not allow to determine separately the partiton coefficients of the protonated and dissociated solute (K_{AH} and K_{A^-}, respectively). It is assumed that the partition coefficient K_T that accounts for the partitioning of both neutral and dissociated forms of the solute is measured (see chapter 4.5). Hence, the partition coefficients of isovanillin at pH = 10.5 are used from reference 42 unmodified (further referred as ionized^{exp}) and compared to the partition coefficients calculated with equation 4.5, chapter 4.5 (further referred as ionized^{calc}). The calculated and experimentally determined partition coefficients for the protonated form of isovanillin are shown for comparison in Figure 7.5. It can be seen that the calculated trends are in

qualitative agreement with the experimental ones. For the dissociated state the absolute deviation from the experimental values is within ∼1 log unit at all molar ratios of CTAB, which is comparable to the deviation of the neutral form for $\alpha \geq 0.8$. Please note that the direct calculation of ionized molecules in mixed micelles with COSMOmic or COSMO-RS without taking into account the protolytic equilibrium within the micelle leads to very high deviations from the experimental values (up to some orders of magnitude), especially in micelles with high nonionic surfactant content. It has to be noted that the partition coefficient of dissociated isovanillin in pure CTAB has an absolute deviation of 0.69 log units, which is in the range of the experimental error. Since this value is used for the calculation of the partition coefficients at $\alpha = 0.2$, $\alpha = 0.5$ and $\alpha = 0.8$, its error contributes to the error of all calculated partition coefficients. If using a partition coefficient for pure CTAB, which is closer to the experimental value, the deviations at the other molar ratios would decrease as well. If using the experimental value in pure CTAB instead of the predicted one for the calculations at the other three molar ratios, the overall RMSE for the four compositions decreases from 0.81 to 0.51, respectively. Moreover, the difference between the calculated and experimental partition coefficient of neutral isovanillin in pure CTAB micelles is ∼1 log unit as well. Hence, the higher deviations in this system can be attributed to the interactions between isovanillin and CTAB in general rather than to the charge of the solute. The same trend was observed for five other solutes in pure CTAB micelles (see chapter 6.3.4).

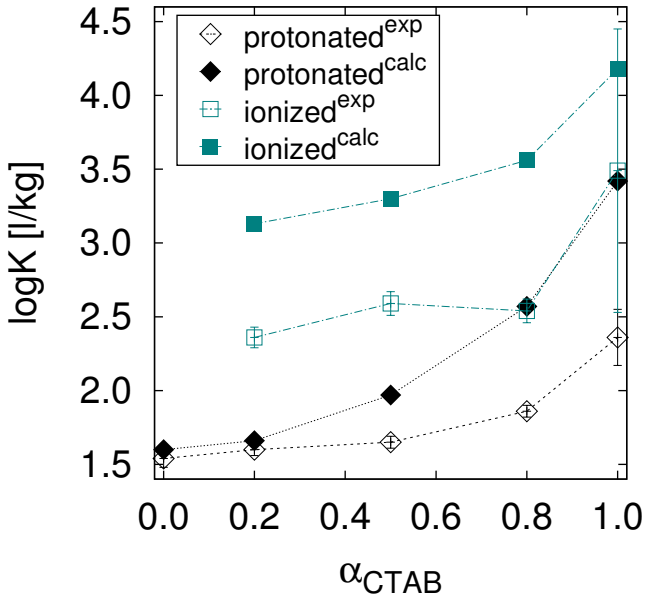

Figure 7.5: Partition coefficients of isovanillin in Brij35/CTAB mixed micelles at different micelle compositions. The experimental data is taken from reference 42. Lines are added as guide to the eye.

Table 7.4: Calculated and experimentally determined partition coefficients of ionized isovanillin in Brij35/CTAB mixed micelles at different micelle compositions, where α is the mole fraction of CTAB, ionized^{calc} is the partition coefficient calculated with COSMOmic using the procedure described in chapter 4.5. The experimental data is taken from reference 42. All K_i have units of L/kg.

α	$\log K_i$	
	ionized^{calc}	ionized^{exp}
0.2	3.13	2.36 ± 0.07
0.5	3.30	2.59 ± 0.08
0.8	3.56	2.54 ± 0.08
1.0	4.18	3.49 ± 0.96

It was demonstrated that it is crucial to take into account the protolytic equilibrium within the micelle for the calculation of partition equilibria of ionized solutes with COSMOmic. By applying a thermodynamic cycle based on transfer free energies calculated with COSMOmic, the estimated partition coefficients are in qualitative agreement with the experimental data and have an absolute deviation within ∼1 log unit. For comparison, the direct calculation with COSMOmic considering only the dissociated form of isovanillin can lead to absolute deviations up to ∼15 log units for the micelles with high content of nonionic surfactant.

7.3 Summary

The aggregation behavior and the composition of mixed TX114/SDS and TX114/CTAB micelles were estimated from self-assembly MD simulations. The composition in the micelles approaches the composition in the initial configuration, which is in accordance with experimental data in the relevant surfactant concentration range.[94,213] Therefore, for the predictions of partition coefficients it was assumed that the composition in the mixed micelle is equal to the composition in the mixed surfactant solution. Both COSMO-RS and COSMOmic provide qualitatively correct predictions in the investigated systems. Nevertheless, the usage of COSMOmic shows advantages over the COSMO-RS pseudo phase approach, as the predictions are generally in better agreement with experimental data. In many cases the COSMOmic calculations result in good quantitative agreement (e.g., see Table 7.1). However, a first condition is that the partition coefficients should be correctly predicted for the pure micelles of the involved surfactants. If this is not the case, micelles with a high content of the inaccurately described surfactant cannot be correctly described (e.g., see

Figure 7.2a). Furthermore, partition behavior of dissociated isovanillin in mixed Brij35/CTAB at different molar ratios of CTAB was investigated. The usage of a thermodynamic cycle based on transfer free energies allows the estimation of the protolytic equilibrium not only in water but also in the micelle. It could be concluded that taking into account the protolytic equilibrium within the micelle is crucial for a reasonable description of partition equilibria of ionized solutes in mixed micelles with COSMOmic. If considering only the ionized form of the solutes, the COSMOmic calculations in mixed micelles lead to much higher deviations from experimental data compared to neutral solutes.

8 Micelle/water compared to liposome/water partition coefficients

In the previous chapters 6 and 7, the prediction quality of the COSMOmic model was evaluated for the partitioning of neutral and dissociated solutes in all micelle types (nonionic, zwitterionic, anionic and caionic) and in mixed nonionic/ionic micelles, whereas in general reliable accuracy was obtained. Zwitterionic surfactants are of special relevance for pharmaceutical studies, as they can be used to mimic the lipophilicity of drugs in biomembranes. The zwitterionic surfactant HePC (miltefosine) is often used as model for biomembranes due to its structural similarity to the phospholipids present in biomembranes and the easier preparation of micelles compared to liposomes.[26–28] It has been demonstrated that the COSMOmic model shows a reliable prediction quality for membrane/water partition coefficients of neutral molecules.[161,194,205,225,240] However, predicted membrane/water and micelle/water partition coefficients obtained with
COSMOmic were not yet critically compared. The resuts of the following section were partly published as Loureiro, D.; Soares, J.; Lopes, D.; Macedo, T.; Yordanova, D.; Jakobtorweihen, S.; Nunes, C.; Reis, S.; Pinto, M.; Afonso, C. **Accessing lipophilicity of drugs with biomimetic models: a comparative study using liposomes and micelles** in European Journal of Pharmaceutical Sciences.[286]

8.1 Simulation details

Averaged atomic distributions of the membrane and micelle are calculated from MD simulations. The DMPC bilayer structure used for the COSMOmic calculations was obtained from a simulation at 303 K with 128 lipid molecules and ionic strength of I=0.1 mol/L. The preassembled HePC micelle containing 100 monomers was simulated at 298 K, with an ionic strength of I=0.1 mol/L. Representative lipid and surfactant conformers were selected from MD simulation with respect to the solvent accessible surface.[240] For all solute molecules the conformers were created with COSMOconf version 3.0.[279] The DFT/COSMO calculations for all molecules, which are required for COSMOmic, were then performed with Turbomole 6.6.[278] The COSMOtherm[280,281] (version C3.0 Release 17.01) implementation of COSMO-RS and COSMOmic with the BP_TZVP_C30_1701 parametrization was used. In the case of charged solutes in DMPC bilayers, the electrostatic potential introduced by Bittermann et al.[171] was used for the calculation.

8.2 Partition coefficients in DMPC/water and HePC/water: COSMOmic versus experiments

In this chapter DMPC/water and HePC/water partition coefficients are predicted and compated to experimental values. DMPC is choosen as a lipid due to the structural similarity with HePC. In Table 8.1 calculated partition coefficients in DMPC lipid bilayer and HePC micelles are given in comparison to the experimental results. The experimental partition coefficients in both liposomes and micelles are determined with derivative spectrophotometry and taken from reference 286. A set of 17 reference compounds is selected from the international accepted "Organisation for Economic Co-operation and Development (OECD) Guidelines for the Testing of Chemicals for the partition coefficient determination - partition coefficient". These compounds constitute frameworks frequently found in approved drugs and/or drugs in different phases of development.[303,304] The experimets are performed at pH = 7.4, at which 2,4-dinitro-6-sec-butylphenol, benzoic acid, phenoxyacetic acid and trans-cinnamic are completely dissociated and 2-nitrophenol is predominantly in the anionic form (86 % dissociated).

Table 8.1: Comparison between experimental and calculated (with COSMOmic) partition coefficients in DMPC and HePC. Experiments and calculations are performed at 310 K. All partition coefficients have units of L/kg.

Solute	DMPC		HePC	
	$\log K_i^{\text{calc}}$	$\log K_i^{\text{exp}}$	$\log K_i^{\text{calc}}$	$\log K_i^{\text{exp}}$
1-Naphthol	3.0	3.4 ± 0.1	2.8	3.4 ± 0.1
2,4-Dinitro-6-sec-butylphenol (anionic)	3.0	3.0 ± 0.1	2.0	3.3 ± 0.1
2,6-Dichlorobenzonitrile	2.7	3.5 ± 0.1	2.4	3.8 ± 0.2
2,6-Diphenylpyridine	5.1	4.1 ± 0.1	4.9	4.6 ± 0.1
2-Nitrophenol (86 % anionic)	2.2	3.3 ± 0.1	1.9	3.8 ± 0.1
4-Phenylphenol	3.4	3.7 ± 0.2	3.2	3.8 ± 0.1
Anisole	2.7	3.6 ± 0.1	2.5	4.1 ± 0.1
Benzoic acid (anionic)	1.5	2.3 ± 0.1	1.1	-
Benzophenone	3.1	3.2 ± 0.1	2.9	3.6 ± 0.1
Diphenyl ether	4.3	3.9 ± 0.1	4.1	4.5 ± 0.1
Naphthalene	3.4	4.4 ± 0.1	3.3	4.8 ± 0.1
Nitrobenzene	2.5	2.6 ± 0.1	2.4	-
p-Cresol	2.4	2.9 ± 0.1	2.3	-
Phenol	2.0	2.8 ± 0.1	1.9	-
Phenoxyacetic acid (anionic)	1.4	2.8 ± 0.2	1.1	-
Thymol	3.8	3.3 ± 0.1	3.6	3.7 ± 0.1
trans-Cinnamic Acid (anionic)	1.6	3.1 ± 0.1	1.3	-
$\mathbf{RMSE}_{\text{neutral}}$	0.65		0.95	
$\mathbf{RMSE}_{\text{anionic}}$	1.11		1.62	
$\mathbf{RMSE}_{\text{all}}$	0.81		1.10	

The COSMOmic predictions show a good agreement to the experimental values in DMPC for the majority of the substances. In fact, the difference between the experimental and the calculated $\log K_i$ values in DMPC are ≤ 0.3 log units for 1-naphthol, 2,4-dinitro-6-sec-butylphenol, 4-phenylphenol, benzophenone and nitrobenzene. This difference is within the experimental error range.[282] The substances could be divided into two groups: neutral substances and dissociated substances at pH 7.4. In the case of neutral substances, higher deviations (ca. 1 log unit) are obtained for 2,6-diphenylpiridine, anisole and naphthalene. Deviations of the COSMOmic predictions for polycyclic aromatic hydrocarbons have been already reported and discussed in the literature.[240] The overall RMSE for ionized compounds (1.11) is significant higher than for neutral

solutes (0.65). Actually, only one of the experimental partition coefficients of the dissociated molecules (2,4-dinitro-6-sec-butylphenol) can be reproduced correctly with COSMOmic. The absolute deviation for the other charged solutes is higher than 1 log unit, whereas the calculations underestimate the experimental findings. Bittermann et al.[171] reported a RMSE of 0.66 for 56 organic anions in phospholipid bilayers. However, these authors used a 0.32 offset, which is not subtracted in the calculations in this work. Taking into account the fact that only five ionized solutes are compared, the prediction quality can be considered as comparable to the findings of Bittermann et al.[171]. The predicted $\log K_i$ values in HePC micelles using COSMOmic, show higher deviations from experimental values than those in DMPC bilayers (RMSE = 0.95 for neutral and 1.62 for ionized molecules, respectively). Please note, that a smaller set of partition coefficients is used for the calculation of RMSE in HePC micelles (only 9 neutral and 2 charged solutes). Experimentally it was obtained that HePC/water partition coefficients are in general higher than for DMPC/water. This trend could not be found by the calculations with COSMOmic. The experimentally found trend was related to the lower membrane fluidity and the rigid structure order of the liposome bilayers compared to the monolayer micelles.[286] However, the prediction of this effect with COSMOmic is difficult, as the density of the system is not accounted for in the model. In contrast to the experimental observations, the predicted values in HePC are slightly lower than in DMPC ($\leq$ 0.3 log units for all solutes except 2,4-dinitro-6-sec-butylphenol). Free energy profiles are examined to gain information about the partition behavior of a solute as a function of its position within the membrane or micelle. In Figure 8.1 the free energy profiles of ionized 2,4-dinitro-6-sec-butylphenol in a DMPC bilayer and in a HePC micelle are shown.

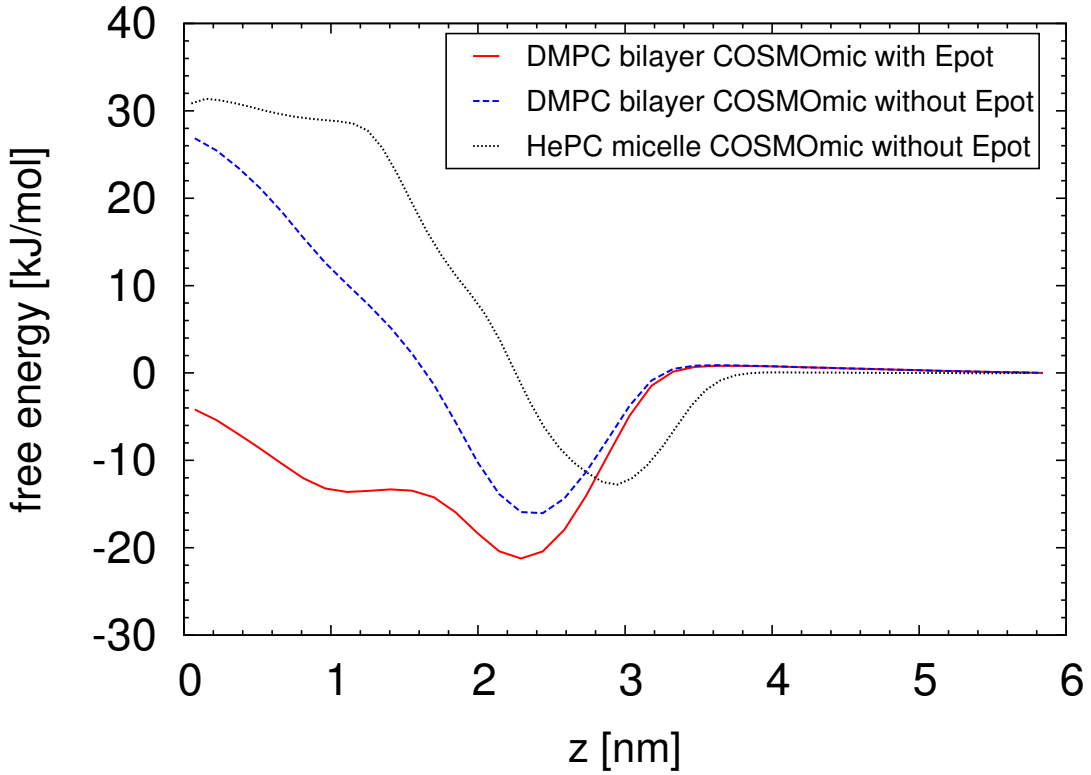

Figure 8.1: Free energy profiles of 2,4-dinitro-6-sec-butylphenol in a HePC micelle and in a DMPC membrane calculated with COSMOmic with and without using the electrostatic potential (Epot) implemented by Bittermann et al.[171] The x-axis represents the distance from the micelle or membrane center.

The profiles in the micelle and in the liposome without electrostatic potential are similar, leading to the prediction of almost identical partition coefficients ($\log K_i$ in HePC of 2.0 and $\log K_i$ in DMPC of 2.1, value not presented in Table 8.1). In contrast, the profile with electrostatic potential shows significant differences, especially in the hydrophobic region (0-2 nm), leading to 1 log unit higher partition coefficient, which is in agreement with the experimental value (see Table 8.1). 2,4-Dinitro-6-sec-butylphenol consists of both an ionizable group interacting with the phosphocholine group and a lipophilic part partitioning in the hydrophobic core. Therefore, the implementation of an internal electrostatic membrane potential is crucial for the reasonable prediction with COSMOmic. Nevertheless, deviations between free energy profiles in the HePC micelle and in the DMPC membrane are obtained for all neutral and charged molecules (data not shown). Mostly, enhanced electrostatic interactions with the hydrophilic group of DMPC are predicted compared to those with the head group of HePC. These can be attributed to the polar glycerol group present in DMPC resulting in higher partition coefficients in DMPC than in HePC.

8.3 Summary

In summary, COSMOmic is able to predict DMPC/water partition coefficients of neutral solutes with sufficient accuracy for screening studies (RMSE = 0.65). A special caution should be taken in case of charged molecules, as in general underestimated values are obtained with a higher deviation from experimental data (RMSE = 1.11). However, the experimentally observed trend that partition coefficients in DMPC/water have in general lower values than those in HePC/water could not be predicted by COSMOmic. In constrast, the opposite trend is obtained, which reflects also in higher quantitative deviations for the system HePC/water. In this case, it can be recommended to use another method (e.g., MD simulations), which takes into account the density of the system and the dynamics of the surfactant aggregates in presence of partitioning solutes.

9 Conclusions

The combination between all-atom MD simulations and the COSMOmic model was evaluated for partition equilibria predictions in various micellar systems including nonionic, zwitterionic, anionic and cationic micelles, as well as nonionic/anionic and nonionic/cationic mixed micelles. In contrast to the most studies on partition equilibria predictions in micellar systems available in the literature, which apply the pseudo phase approach the methods used in this work consider the anisotropic structure of micelles. In zwitterionic, cationic and nonionic/cationic micelles, negatively charged solutes were studied with a special focus. In this work, the prediction quality of the combination between MD simulations and COSMOmic for charged solutes was evaluated for the first time. The predictions were validated based on comparison with experimentally determined partition coefficients. Moreover, to gain a detailed information on the localization of solutes within the micelles, free energy profiles were obtained, where two approaches were compared: MD simulations with the umbrella sampling (US) method and COSMOmic.

The nonionic surfactants Triton X-114 and Triton X-100 were of particular interest in this work due to their wide range of practical applications. Since the Triton X molecules were not present in any known biomolecular force field, the first optimized set of Triton X parameters for the CHARMM general force field has been developed. Therefore, the self-assembly of Triton X surfactants could be studied for the first time via all-atom MD simulations. Micelle structures obtained from the MD simulations were used as input for the model COSMOmic in order to predict the partition behavior of neutral solutes. In order to achieve statistical reliable results, averaged atomic distributions for the micelle structure were used. This approach reduces the effect of outliers. By comparing the predictions of the partition coefficients using different micelle structures, an influence of the micelle size and shape on the prediction quality was observed. When using large or more cylindrical micelles, outliers are more likely to occur. Whereas the micelle shape is problematic for large micelles, for small micelles the water layer around the micelle can be problematic. In self-assembly simulations other micelles or surfactant molecules are present such that sometimes no bulk water is found around the micelle. This problem was solved in this work by replacing the disturbing molecules by water in the distributions used for COSMOmic.

Free energy profiles calculated by applying two approaches were firstly compared: MD simulations with the US method and the COSMOmic model. Although some deviations were observed, the free energy profiles of neutral solutes

obtained with both methods are in an overall good agreement. Additionally, negatively charged compounds were studied in zwitterionic HePC and cationic CTAB micelles. To make a quantitative comparison possible, partition coefficients of dissociated solutes in CTAB micelles were determined using micellar liquid chromatography (MLC). It could be demonstrated that for charged compounds, the experimental determination of partition equilibria with the MLC method can be challenging due to the so-called overbinding of the solutes. In case of a dissociated acid in a HePC micelle, a decrease in the partition coefficient of the dissociated acid compared to its neutral form is expected. This behavior could be predicted qualitatively correct with both models. In case of negatively charged solutes in CTAB micelles, the opposite behavior is expected: the partition coefficient of the dissociated states should increase compared to the neutral form due to enhanced electrostatic interactions between the negatively charged solute and the positively charged head group of CTAB. This trend was not obtained with MD simulations. It is expected that the usage of polarizable force fields for this system could be advantageous. In contrast to MD simulations, COSMOmic is able to predict a favorable partitioning of dissociated solutes in CTAB. Nevertheless, the COSMOmic calculations for charged molecules should be interpreted carefully. It is known that COSMOmic shows higher deviations for predicted partition coefficients of charged solutes in lipid bilayers compared to those of neutral. Bittermann et al.[171] improved the prediction quality for charged molecules in DMPC membranes by the implementation of an internal membrane dipole potential, which was empirically optimized by using experimental partition coefficients. However, the introduction of a dipole potential for a micelle for COSMOmic analog to lipid bilayers is only possible, if a large set of partition coefficients for charged solutes in one type of micelle is available. Unfortunately, experimental data for charge compounds in micelles are scarce in the literature.

The combination between MD simulations and COSMOmic was further applied to study the micellization and partition behavior in various mixed surfactant systems. The COSMO-RS pseudo phase approach and COSMOmic were compared for the calculation of partition coefficients of neutral solutes in mixed nonionic/ionic micelles, whereby both provide qualitatively correct predictions. Nevertheless, the usage of COSMOmic shows advantages over COSMO-RS, as the predictions are generally in better agreement with experimental data. Furthermore, a thermodynamic cycle based on transfer free energies was applied to obtain a position dependent pK_{a} value in the micelles. This approach allows the estimation of the protolytic equilibrium within the micelle and was used for partition coefficient predictions of dissociated isovanillin in mixed Brij35/CTAB micelles. For this system, taking into account the protolytic equilibrium within

the micelle was crucial for a reasonable description of partition equilibria with COSMOmic. If considering only the ionized form of the solute, the COSMOmic calculations in mixed micelles lead to much higher deviations from experimental data compared to neutral solutes. It can be concluded that COSMOmic can qualitatively predict the influence of surfactant mixtures on partition behavior of solutes. However, a first condition is that the partition coefficients should be correctly predicted for the pure micelles of the involved surfactants. If this is not the case, micelles with a high content of the inaccurately described surfactant cannot be correctly described.

In summary, in this work the combination of MD simulations with the COSMOmic model was successfully applied for the prediction of partition coefficients in single and mixed micellar systems, whereby an overall good prediction quality for neutral solutes was obtained. Criteria to select suitable micelle structures for the COSMOmic calculations were proposed and techniques to reduce the effect of outliers were established. Additionaly, free energy profiles of neutral and charged solutes in various types of micelles were obtained using two different methods. Finally, an approach, which takes into account the protolytic equilibrium within the micelle was proposed and applied for the partition coefficients prediction of negatively charged solute in mixed nonionic/cationic micelles. By applying the proposed approach, a significant improvement in the prediction quality of COSMOmic could be achieved.

Bibliography

[1] Birdi, K. S., Ed. *Handbook of surface and colloid chemistry*, fourth edition ed.; CRC Press, Taylor & Francis Group: Boca Raton, 2015.

[2] Williams, R. O. B. *J. Am. Chem. Soc.* **2000**, *122*, 9882–9882.

[3] Christian, S. D., Scamehorn, J. F., Eds. *Solubilization in surfactant aggregates*; Surfactant science series v. 55; M. Dekker: New York, 1995.

[4] Solans, C., Kunieda, H., Eds. *Industrial applications of microemulsions*; Surfactant science series v. 66; M. Dekker: New York, 1997.

[5] Uddin, M. H.; Kanei, N.; Kunieda, H. *Langmuir* **2000**, *16*, 6891–6897.

[6] Tokuoka, Y.; Uchiyama, H.; Abe, M.; Christian, S. D. *Langmuir* **1995**, *11*, 725–729.

[7] Tokuoka, Y.; Uchiyama, H.; Abe, M.; Ogino, K. *J. Colloid Interface Sci.* **1992**, *152*, 402–409.

[8] Helenius, A.; Simons, K. *Biochim. Biophys. Acta* **1975**, *415*, 29–79.

[9] Tanford, C.; Reynolds, J. A. *Biochim. Biophys. Acta* **1976**, *457*, 133–170.

[10] Wennerström, H. *Phys. Rep.* **1979**, *52*, 1–86.

[11] Lichtenberg, D.; Robson, R. J.; Dennis, E. A. *Biochim. Biophys. Acta* **1983**, *737*, 285–304.

[12] le Maire, M.; Champeil, P.; Møller, J. V. *Biochim. Biophys. Acta* **2000**, *1508*, 86–111.

[13] Dickie, P.; Weiner, J. H. *Canadian J. Biochem.* **1979**, *57*, 813–821.

[14] Kataoka, K.; Harada, A.; Nagasaki, Y. *Adv. Drug Deliver Rev.* **2001**, *47*, 113–131.

[15] Nasongkla, N.; Bey, E.; Ren, J.; Ai, H.; Khemtong, C.; Guthi, J. S.; Chin, S.-F.; Sherry, A. D.; Boothman, D. A.; Gao, J. *Nano Lett.* **2006**, *6*, 2427–2430.

[16] Duester, L.; Fabricius, A.-L.; Jakobtorweihen, S.; Philippe, A.; Weigl, F.; Wimmer, A.; Schuster, M.; Nazar, M. F. *Anal. Bioanal. Chem.* **2016**, *408*, 7551–7557.

[17] Chen, J.; Teo, K. C. *Anal. Chim. Acta.* **2001**, *434*, 325–330.

[18] Tang, A. N.; Jiang, D. Q.; Yan, X. P. *Anal. Chim. Acta.* **2004**, *507*, 199–204.

[19] Niazi, A.; Momeni-Isfahani, T.; Ahmari, Z. *J. Hazard. Mater.* **2009**, *165*, 1200–1203.

[20] Arunagiri, A.; Priya, K.; Kalaichelvi, P.; Anantharaj, R. *J. Ind. Eng. Chem.* **2014**, *20*, 2409–2420.

[21] Xing, W.; Chen, L.; Zhang, F. *Anal. Methods* **2014**, *6*, 3644.

[22] Glembin, P.; Racheva, R.; Kerner, M.; Smirnova, I. *Sep. Purif. Technol.*

2014, *135*, 127–134.

[23] Ingram, T.; Storm, S.; Glembin, P.; Bendt, S.; Huber, D.; Mehling, T.; Smirnova, I. *Chem. Ing. Tech.* **2012**, *84*, 840–848.

[24] Ritter, E.; Racheva, R.; Jakobtorweihen, S.; Smirnova, I. *Chem. Eng. Res. Des.* **2017**,

[25] Xia, Q.; Jiao, Y.; Xiong, W.; Yang, Y.; Liu, M. *Food Anal. Methods* **2014**, *7*, 1130–1138.

[26] Magalhães, L. M.; Nunes, C.; Lúcio, M.; Segundo, M. A.; Reis, S.; Lima, J. L. F. C. *Nat. Protoc.* **2010**, *5*, 1823–1830.

[27] De Vrieze, M.; Janssens, P.; Szucs, R.; Van der Eycken, J.; Lynen, F. *Anal. Bioanal. Chem.* **2015**, *407*, 7453–7466.

[28] de Castro, B.; Gameiro, P.; Lima, J. L.; Matos, C.; Reis, S. *Mat. Sci. Eng. C* **2001**, *18*, 71–78.

[29] Hill, R. In *Mixed Surfactant Systems*; Ogino, A. M., K., Ed.; M. Dekker: New York: New York, 1993; pp 317–336.

[30] Abe, M.; Ogino, K. In *Mixed Surfactant Systems*; Ogino, A. M., K., Ed.; M. Dekker: New York: New York, 1993; pp 1–21.

[31] Hines, J. D. *Curr. Opin. Colloid Interface Sci.* **2001**, *6*, 350–356.

[32] Ebrahim Attia, A. B.; Ong, Z. Y.; Hedrick, J. L.; Lee, P. P.; Ee, P. L. R.; Hammond, P. T.; Yang, Y.-Y. *Curr. Opin. Colloid Interface Sci.* **2011**, *16*, 182–194.

[33] Gharibi, H.; Razavizadeh, B.; Hashemianzaheh, M. *Colloids Surf. A* **2000**, *174*, 375–386.

[34] Szymczyk, K.; Zdziennicka, A.; Krawczyk, J.; Jaczuk, B. *J. Chem. Thermodyn.* **2014**, *69*, 85–92.

[35] Austin, R. P.; Davis, A. M.; Manners, C. N. *J. Pharm. Sci.* **1995**, *84*, 1180–1183.

[36] Avdeef, A. *Absorption and drug development: solubility, permeability, and charge state*, 2nd ed.; John Wiley & Sons: Hoboken, N.J, 2012.

[37] Avdeef, A.; Box, K. J.; Comer, J. E. A.; Hibbert, C.; Tam, K. Y. *Pharm. Res.* **1998**, *15*, 209–215.

[38] Bhat, P. A.; Rather, G. M.; Dar, A. A. *J. Phys. Chem. B* **2009**, *113*, 997–1006.

[39] Rupp, C.; Steckel, H.; Mller, B. W. *Int. J. Pharm.* **2010**, *395*, 272–280.

[40] Schwebel, H. J.; van Hoogevest, P.; Leigh, M. L.; Kuentz, M. *Pharm. Dev. Technol.* **2011**, *16*, 278–286.

[41] Mehling, T.; Kloss, L.; Ingram, T.; Smirnova, I. *Langmuir* **2013**, *29*, 1035–1044.

[42] Mehling, T. Partition coefficients in mixed surfactant systems. PhD thesis, Technische Universitt Hamburg-Harburg, 2013.

[43] Materna, K.; Goralska, E.; Sobczynska, A.; Szymanowski, J. *Green Chem.* **2004**, *6*, 176.
[44] Kadam, Y.; Yerramilli, U.; Bahadur, A. *Colloids Surf., B* **2009**, *72*, 141–147.
[45] Treiner, C.; Kumar Chattopadhyay, A. *J. Colloid Interface Sci.* **1986**, *109*, 101–108.
[46] Taft, R. W.; Abboud, J.-L. M.; Kamlet, M. J.; Abraham, M. H. *J. Solution Chem.* **1985**, *14*, 153–186.
[47] Hansch, C.,; Leo, A., *Fundamentals and Applications in Chemistry and Biology, ACS Professional Reference Book*; American Chemical Society (ACS): Washington, DC, 1995.
[48] Hu, J.; Zhang, X.; Wang, Z. *Int. J. Mol. Sci.* **2010**, *11*, 1020–1047.
[49] Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224–2235.
[50] Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.
[51] Blankschtein, D.; Thurston, G. M.; Benedek, G. B. *Phys. Rev. Lett.* **1985**, *54*, 955–955.
[52] Klamt, A.; Huniar, U.; Spycher, S.; Keldenich, J. *J. Phys. Chem. B* **2008**, *112*, 12148–12157.
[53] Lasic, D. D. *Trends Biotechnol.* **1998**, *16*, 307–321.
[54] Hill, J.; Shrestha, L.; Ishihara, S.; Ji, Q.; Ariga, K. *Molecules* **2014**, *19*, 8589–8609.
[55] Epand, R. M.; D'Souza, K.; Berno, B.; Schlame, M. *BBA - Biomembranes* **2015**, *1848*, 220–228.
[56] Lasic, D. D. *Liposomes: from physics to applications*; Elsevier: Amsterdam ; New York, 1993.
[57] Rangel-Yagui, C.; Pessoa, A. J.; Tavares, L. *J. Pharm. Pharmaceut. Sci.* **2005**, *8*, 147–163.
[58] Attwood, D.; Florence, A. T. *Physical pharmacy*, 2nd ed.; FASTtrack; Pharmaceutical Press: London, 2012; OCLC: 935057580.
[59] Puvvada, S.; Blankschtein, D. *J. Chem. Phys.* **1990**, *92*, 3710–3724.
[60] Rosen, M. J.; Kunjappu, J. T. *Surfactants and interfacial phenomena*, 4th ed.; Wiley: Hoboken, N.J, 2012.
[61] Hinze, W. L.; Pramauro, E. *Crit. Rev. Anal. Chem.* **1993**, *24*, 133–177.
[62] Attwood, D.; Florence, A. T. *Surfactant Systems: Their chemistry, pharmacy and biology*; Springer Netherlands: Dordrecht, 1983; OCLC: 858945678.
[63] Alvarez-Nez, F.; Yalkowsky, S. *Int. J. Pharm.* **2000**, *200*, 217–222.
[64] Blazyk, J. F.; Steim, J. M. *BBA - Biomembranes* **1972**, *266*, 737–741.
[65] Eze, M. O. *Biochem. Educ.* **1991**, *19*, 204–208.

[66] Seddon, J. M. *Biochim. Biophys. Acta* **1990**, *1031*, 1–69.
[67] Yatvin, M. B.; Kreutz, W.; Horwitz, B. A.; Shinitzky, M. *Science* **1980**, *210*, 1253–1255.
[68] Hafez, I. M.; Ansell, S.; Cullis, P. R. *Biophys. J.* **2000**, *79*, 1438–1446.
[69] Myers, D. *Surfactant Science and Technology: Myers/Surfactant Science and Technology, 3rd Edition*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2005; DOI: 10.1002/047174607X.
[70] Tadros, T. F. *Applied surfactants: principles and applications*, 2nd ed.; Wiley-VCH: Weinheim, 2008; OCLC: 699808070.
[71] Walker, C. *FEBS Letters* **1981**, *124*, 127–127.
[72] Chandler, D. *Nature* **2005**, *437*, 640–647.
[73] Kile, D. E.; Chiou, C. T. *Environ. Sci. Technol.* **1989**, *23*, 832–838.
[74] Mukerjee, P. *J. Phys. Chem.* **1962**, *66*, 1375–1376.
[75] Kamrath, R. F.; Franses, E. I. *J. Phys. Chem.* **1984**, *88*, 1642–1648.
[76] Jakobtorweihen, S.; Yordanova, D.; Smirnova, I. *Chem. Ing. Tech.* **2017**, *89*, 1288–1296.
[77] Jones, M. N.; Chapman, D. *Micelles, monolayers, and biomembranes*; Wiley-Liss: New York, 1995.
[78] Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. *J. Chem. Soc., Faraday Trans. 2* **1976**, *72*, 1525.
[79] Tanford, C. *J. Phys. Chem.* **1974**, *78*, 2469–2479.
[80] Nagarajan, R.; Ruckenstein, E. *Langmuir* **1991**, *7*, 2934–2969.
[81] Nagarajan, R. *Langmuir* **2002**, *18*, 31–38.
[82] Huibers, P. D. T.; Lobanov, V. S.; Katritzky, A. R.; Shah, D. O.; Karelson, M. *Langmuir* **1996**, *12*, 1462–1470.
[83] Huibers, P.; Lobanov, V.; Katritzky, A.; Shah, D.; Karelson, M. *J. Colloid Interface Sci.* **1997**, *187*, 113–120.
[84] Jalali-Heravi, M.; Konouz, E. *J. Surfactants Deterg.* **2000**, *3*, 47–52.
[85] Kardanpour, Z.; Hemmateenejad, B.; Khayamian, T. *Anal. Chim. Acta* **2005**, *531*, 285–291.
[86] Yuan, S.; Cai, Z.; Xu, G.; Jiang, Y. *Colloid Polym. Sci.* **2002**, *280*, 630–636.
[87] Chen, C.-C. *AIChE J.* **1996**, *42*, 3231–3240.
[88] Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. *AIChE Journal* **1975**, *21*, 1086–1099.
[89] Li, X.-S.; Lu, J.-F.; Li, Y.-G.; Liu, J.-C. *Fluid Phase Equilibr.* **1998**, *153*, 215–229.
[90] Li, X.-S.; Lu, J.-F.; Li, Y.-G. *Fluid Phase Equilibr.* **2000**, *168*, 107–123.
[91] Holland, P. M.; Rubingh, D. N. In *Mixed Surfactant Systems*; Holland, P. M., Rubingh, D. N., Eds.; American Chemical Society: Washing-

ton, DC, 1992; Vol. 501; pp 2–30, DOI: 10.1021/bk-1992-0501.ch001.
[92] Glenn, K.; van Bommel, A.; Bhattacharya, S. C.; Palepu, R. M. *Colloid Polym. Sci.* **2005**, *283*, 845–853.
[93] Clint, J. H. *J. Chem. Soc., Faraday Trans. 1* **1975**, *71*, 1327.
[94] Gao, H.; Zhu, R.; Yang, X.; Mao, S.; Zhao, S.; Yu, J.; Du, Y. *J. Colloid Interface Sci.* **2004**, *273*, 626–631.
[95] Rubingh, D. N. In *Solution Chemistry of Surfactants*; Mittal, K. L., Ed.; Springer New York: Boston, MA, 1979; pp 337–354.
[96] Shinoda, K. *J. Phys. Chem.* **1954**, *58*, 541–544.
[97] Christov, M. *Thermodynamik*; Vanio Nedkov Verlag: Universität für chemische Technologie und Metallurgie Sofia, 2003.
[98] Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525–616.
[99] Ingram, T.; Richter, U.; Mehling, T.; Smirnova, I. *Fluid Phase Equilibr.* **2011**, *305*, 197–203.
[100] Marina, M. L.; Garca, M. A. *Encyclopedia of Separation Science*; Academic Press Ltd: England, 2000; Vol. II; pp 726–737.
[101] Armstrong, D. W.; Nome, F. *Anal. Chem.* **1981**, *53*, 1662–1666.
[102] M. Basova, E.; M. Ivanov, V.; A. Shpigun, O. *Russ. Chem. Rev.* **1999**, *68*, 983–1000.
[103] Sun, J.; Mao, J.; Liu, X.; Wang, Y.; Sun, Y.; He, Z. *J. Sep. Sci.* **2009**, *32*, 2043–2050.
[104] Arunyanart, M.; Love, L. J. C. *Anal. Chem.* **1984**, *56*, 1557–1561.
[105] Mehling, T.; Kloss, L.; Mushardt, H.; Ingram, T.; Smirnova, I. *J. Chromatogr. A* **2013**, *1273*, 66–72.
[106] Jensen, J.-H. Anwendung und Optimierung von MD Simulationen und COSMOmic Berechnungen zur Vorhersage des Verteilungsverhaltens ionischer und nichtionischer Solute in Mizellen. Master's thesis, Technische Universitt Hamburg-Harburg, 2016.
[107] Hill, T. L. *An introduction to statistical thermodynamics*; Dover Publications: New York, 1986.
[108] Gibbs, J. W. *Elementary Principles in Statistical Mechanics*; Charles Scribner's Sons: New York, 1902.
[109] Tuckerman, M. E. *Statistical mechanics: theory and molecular simulation*; Oxford University Press: Oxford ; New York, 2010; OCLC: ocn551495372.
[110] Leach, A. R. *Molecular modelling: principles and applications*, 2nd ed.; Prentice Hall: Harlow, England ; New York, 2001.
[111] Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86–104.
[112] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2009**, *31*, 671–690.

[113] Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
[114] Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
[115] Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2009**, *31*, 1569–1582.
[116] Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2011**, *40*, 843–856.
[117] Scott, W. R. P.; Hnenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
[118] Chiu, S.-W.; Pandit, S. A.; Scott, H. L.; Jakobsson, E. *J. Phys. Chem. B* **2009**, *113*, 2748–2763.
[119] Kukol, A. *J. Chem. Theory Comput.* **2009**, *5*, 615–626.
[120] Poger, D.; Van Gunsteren, W. F.; Mark, A. E. *J. Comput. Chem.* **2010**, *31*, 1117–1125.
[121] Pastor, R. W.; MacKerell, A. D. *J. Phys. Chem. Lett.* **2011**, *2*, 1526–1532.
[122] MacKerell, A. D. et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
[123] MacKerell, A. D.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257–265.
[124] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
[125] Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. *J. Chem. Theory Comput.* **2014**, *10*, 865–879.
[126] Dickson, C. J.; Rosso, L.; Betz, R. M.; Walker, R. C.; Gould, I. R. *Soft Matter* **2012**, *8*, 9617.
[127] Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
[128] Maciejewski, A.; Pasenkiewicz-Gierula, M.; Cramariuc, O.; Vattulainen, I.; Rog, T. *J. Phys. Chem. B* **2014**, *118*, 4571–4581.
[129] Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
[130] Tang, X.; Koenig, P. H.; Larson, R. G. *J. Phys. Chem. B* **2014**, *118*, 3864–3880.
[131] Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667–683.
[132] Loncharich, R. J.; Brooks, B. R. *Proteins: Struct., Funct., Genet.* **1989**, *6*, 32–45.
[133] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

[134] Petersen, H. G. *J. Chem. Phys.* **1995**, *103*, 3668–3679.
[135] Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370.
[136] Raman, E. P.; Guvench, O.; MacKerell, A. D. *J. Phys. Chem. B* **2010**, *114*, 12981–12994.
[137] Mallajosyula, S. S.; Guvench, O.; Hatcher, E.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 759–776.
[138] Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180.
[139] Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
[140] Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmller, H.; MacKerell, A. D. *Nat. Methods* **2017**, *14*, 71–73.
[141] Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.
[142] Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D. *J. Comput. Chem.* **2011**, *32*, 1929–1943.
[143] Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. *J. Comput. Chem.* **2013**, *34*, 2757–2770.
[144] Vanommeslaeghe, K.; MacKerell, A. D. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.
[145] Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.
[146] Mller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
[147] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
[148] Froese Fischer, C. *Comput. Phys. Commun.* **1987**, *43*, 355–365.
[149] Abraham, M.; Hess, B.; van der Spoel, D.; Lindahl, E. GROMACS User Manual version 5.1.1. 2015.
[150] Straatsma, T. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1988**, *89*, 5876–5886.
[151] Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc* **1987**, *109*, 1607–1614.
[152] Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578–581.
[153] Kästner, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932–942.
[154] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
[155] Kstner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104.
[156] Frenkel, D.; Smit, B. *Understanding molecular simulation: from algo-*

rithms to applications, 2nd ed.; Computational science series 1; Academic Press: San Diego, 2002.

[157] Schmidt, R. K.; Teo, B.; Brady, J. W. *J. Phys. Chem.* **1995**, *99*, 11339–11343.

[158] Souaille, M.; Roux, B. *Computer Physics Communications* **2001**, *135*, 40–57.

[159] Trzesniak, D.; Kunz, A.-P. E.; van Gunsteren, W. F. *Chem. Phys. Chem.* **2007**, *8*, 162–169.

[160] Khavrutskii, I. V.; Dzubiella, J.; McCammon, J. A. *J. Chem. Phys.* **2008**, *128*, 044106.

[161] Jakobtorweihen, S.; Zuniga, A. C.; Ingram, T.; Gerlach, T.; Keil, F. J.; Smirnova, I. *J. Chem. Phys.* **2014**, *141*, 045102.

[162] Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

[163] Klamt, A.; Eckert, F.; Hornig, M.; Beck, M. E.; Bürger, T. *J. Comput. Chem.* **2002**, *23*, 275–281.

[164] Klamt, A.; Eckert, F.; Arlt, W. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101–122.

[165] Klamt, A.; Eckert, F. *Fluid Phase Equilib.* **2000**, *172*, 43–72.

[166] Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *Fluid Phase Equilib.* **1998**, *285*, 5074–5085.

[167] Klamt, A. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 699–709.

[168] Spuhl, O. Anwendung quantenchemischer Methoden in der Thermodynamik der Stoff- trennung. PhD thesis, Friedrich-Alexander-Universitt Erlangen-Nürnberg, 2006.

[169] Klamt, A. *COSMO-RS: from quantum chemistry to fluid phase thermodynamics and drug design*, 1st ed.; Elsevier: Amsterdam, 2005; OCLC: 254610452.

[170] Storm, S.; Jakobtorweihen, S.; Smirnova, I. *J. Phys. Chem. B* **2014**, *118*, 3593–3604.

[171] Bittermann, K.; Spycher, S.; Endo, S.; Pohler, L.; Huniar, U.; Goss, K.-U.; Klamt, A. *J. Phys. Chem. B* **2014**, 141212062113002.

[172] De Nicola, A.; Kawakatsu, T.; Rosano, C.; Celino, M.; Rocco, M.; Milano, G. *J. Chem. Theory Comput.* **2015**, *11*, 4959–4971.

[173] Yordanova, D.; Smirnova, I.; Jakobtorweihen, S. *J. Chem. Theory Comput.* **2015**, *11*, 2329–2340.

[174] Denkova, P. S.; Lokeren, L. V.; Verbruggen, I.; Willem, R. *J. Phys. Chem. B* **2008**, *112*, 10935–10941.

[175] Brown, W.; Rymden, R.; Van Stam, J.; Almgren, M.; Svensk, G. *J. Phys. Chem.* **1989**, *93*, 2512–2519.

[176] Jose, B. J.; Bales, B. L.; Peric, M. *J. Phys. Chem. B* **2009**, *113*, 13257–13262.
[177] Espinosa, Y. R.; Caffarena, E. R.; Martnez, Y. B.; Grigera, J. R. *J. Chem. Phys.* **2018**, *148*, 074901.
[178] Fuchs, P. F. J.; Hansen, H. S.; Hnenberger, P. H.; Horta, B. A. C. *J. Chem. Theory Comput.* **2012**, *8*, 3943–3963.
[179] Marrink, S. J.; Tieleman, D. P.; Mark, A. E. *J. Phys. Chem. B* **2000**, *104*, 12165–12173.
[180] Bruce, C. D.; Berkowitz, M. L.; Perera, L.; Forbes, M. D. E. *J. Phys. Chem. B* **2002**, *106*, 3788–3793.
[181] Sammalkorpi, M.; Karttunen, M.; Haataja, M. *J. Phys. Chem. B* **2007**, *111*, 11722–11733.
[182] Sammalkorpi, M.; Karttunen, M.; Haataja, M. *J. Phys. Chem. B* **2009**, *113*, 5863–5870.
[183] Sanders, S. A.; Sammalkorpi, M.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **2012**, *116*, 2430–2437.
[184] Espinosa Silva, Y. R.; Grigera, J. R. *RSC Adv.* **2015**, *5*, 70005–70009.
[185] Storm, S.; Jakobtorweihen, S.; Smirnova, I.; Panagiotopoulos, A. Z. *Langmuir* **2013**, *29*, 11582–11592.
[186] Anachkov, S. E.; Danov, K. D.; Basheva, E. S.; Kralchevsky, P. A.; Ananthapadmanabhan, K. P. *Adv. Colloid Interface Sci.* **2012**, *183-184*, 55–67.
[187] Padalkar, K. V.; Gaikar, V. G.; Aswal, V. K. *Pramana* **2008**, *71*, 953–957.
[188] Snchez, F.; Ruiz, C. *J. Lumin.* **1996**, *69*, 179–186.
[189] Chen, J.; Hao, J. *Phys. Chem. Chem. Phys.* **2013**, *15*, 5563.
[190] Morrow, B. H.; Koenig, P. H.; Shen, J. K. *Langmuir* **2013**, *29*, 14823–14830.
[191] Cistola, D. P.; Atkinson, D.; Hamilton, J. A.; Small, D. M. *Biochemistry* **1986**, *25*, 2804–2812.
[192] Hargreaves, W. R.; Deamer, D. W. *Biochemistry* **1978**, *17*, 3759–3768.
[193] Cistola, D. P.; Hamilton, J. A.; Jackson, D.; Small, D. M. *Biochemistry* **1988**, *27*, 1881–1888.
[194] Lopes, D.; Jakobtorweihen, S.; Nunes, C.; Sarmento, B.; Reis, S. *Prog. Lipid Res.* **2017**, *65*, 24–44.
[195] Matubayasi, N.; Liang, K. K.; Nakahara, M. *J. Chem. Phys.* **2006**, *124*, 154908.
[196] Fujimoto, K.; Yoshii, N.; Okazaki, S. *J. Chem. Phys.* **2012**, *136*, 014511.
[197] Yan, H.; Cui, P.; Liu, C.-B.; Yuan, S.-L. *Langmuir* **2012**, *28*, 4931–4938.
[198] Yuan, F.; Larson, R. G. *J. Phys. Chem. B* **2015**, *119*, 12540–12551.
[199] Grafmüller, A.; Lipowsky, R.; Knecht, V. *Phys. Chem. Chem. Phys.* **2013**, *15*, 876–881.

[200] Buggert, M.; Mokrushina, L.; Smirnova, I.; Schomcker, R.; Arlt, W. *Chem. Eng. Technol.* **2006**, *29*, 567–573.
[201] Mokrushina, L.; Buggert, M.; Smirnova, I.; Arlt, W.; Schomäcker, R. *Ind. Eng. Chem. Res.* **2007**, *46*, 6501–6509.
[202] Buggert, M.; Cadena, C.; Mokrushina, L.; Smirnova, I.; Maginn, E. J.; Arlt, W. *Chem. Eng. J.* **2009**, *32*, 977–986.
[203] Mokrushina, L.; Yamin, P.; Sponsel, E.; Arlt, W. *Fluid Phase Equilib.* **2012**, *334*, 37–42.
[204] Wille, S.; Mokrushina, L.; Smirnova, I.; Arlt, W. *Chem.Eng. Technol.* **2010**, *33*, 2095–2101.
[205] Ingram, T.; Storm, S.; Kloss, L.; Mehling, T.; Jakobtorweihen, S.; Smirnova, I. *Langmuir* **2013**, *29*, 3527–3537.
[206] Ritter, E.; Yordanova, D.; Gerlach, T.; Smirnova, I.; Jakobtorweihen, S. *Fluid Phase Equilib.* **2016**, *422*, 43–55.
[207] Blankschtein, D.; Shiloach, A.; Zoeller, N. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 294–300.
[208] Thomas, H. G.; Lomakin, A.; Blankschtein, D.; Benedek, G. B. *Langmuir* **1997**, *13*, 209–218.
[209] Penfold, J.; Staples, E.; Thompson, L.; Tucker, I.; Hines, J.; Thomas, R. K.; Lu, J. R.; Warren, N. *J. Phys. Chem. B* **1999**, *103*, 5204–5211.
[210] Pedone, L.; Chillura Martino, D.; Caponetti, E.; Floriano, M. A.; Triolo, R. *J. Phys. Chem. B* **1997**, *101*, 9525–9531.
[211] Brasher, L. L.; Kaler, E. W. *Langmuir* **1996**, *12*, 6270–6276.
[212] Griffiths, P. C.; Stilbs, P.; Paulsen, K.; Howe, A. M.; Pitt, A. R. *J. Phys. Chem. B* **1997**, *101*, 915–918.
[213] Fang, X.; Zhao, S.; Mao, S.; Yu, J.; Du, Y. *Colloid Polym. Sci.* **2003**, *281*, 455–460.
[214] Koneva, A. S.; Ritter, E.; Anufrikov, Y. A.; Lezov, A. A.; Klestova, A. O.; Smirnova, N. A.; Safonova, E. A.; Smirnova, I. *Colloids Surf. A* **2018**, *538*, 45–55.
[215] Escher, B. I.; Schwarzenbach, R. P. *Environ. Sci. Technol.* **1996**, *30*, 260–270.
[216] Escher, B. I.; Schwarzenbach, R. P.; Westall, J. C. *Environ. Sci. Technol.* **2000**, *34*, 3954–3961.
[217] Avdeef, A.; Box, K. J.; Comer, J. E. A.; Hibbert, C.; Tam, K. Y. *Pharm. Res.* **1988**, *15*, 209–215.
[218] Miyazaki, J.; Hideg, K.; Marsh, D. *Biochim. Biophys. Acta* **1992**, *1103*, 62–68.
[219] Pauletti, G. M.; Wunderli-Allenspach, H. *Eur. J. Pharm. Sci.* **1994**, *1*,

273–282.
[220] Ottiger, C.; Wunderli-Allenspach, H. *Eur. J. Pharm. Sci.* **1997**, *5*, 223–231.
[221] Čudina, O.; Brborić, J.; Janković, I.; Karljiković-Rajić, K.; Vladimirov, S. *Colloids Surf. B* **2008**, *65*, 80–84.
[222] Flewelling, R.; Hubbell, W. *Biophys. J.* **1986**, *49*, 541–552.
[223] Franklin, J.; Cafiso, D. *Biophys. J.* **1993**, *65*, 289–299.
[224] Wang, L. *Annu. Rev. Biochem.* **2012**, *81*, 615–635.
[225] Bittermann, K.; Spycher, S.; Goss, K.-U. *Chemosphere* **2016**, *144*, 382–391.
[226] Schreier, S.; Frezzatti, W. A.; Araujo, P. S.; Chaimovich, H.; Cuccovia, I. M. *BBA - Biomembranes* **1984**, *769*, 231–237.
[227] Beschiaschvili, G.; Seelig, J. *Biochem.* **1992**, *31*, 10044–10053.
[228] Rodgers, A. H.; Khaledi, M. G. *Anal. Chem.* **1994**, *66*, 327–334.
[229] Mchedlov-Petrossyan, N. O. *Pure Appl. Chem.* **2008**, *80*.
[230] Alongi, K. S.; Shields, G. C. *Annual Reports in Computational Chemistry*; Elsevier, 2010; Vol. 6; pp 113–138.
[231] Li, G.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 14521–14528.
[232] Pliego, J. R. *Chem. Phys. Lett.* **2003**, *367*, 145–149.
[233] Jover, J.; Bosque, R.; Sales, J. *QSAR Comb. Sci.* **2008**, *27*, 563–581.
[234] Chew, C. F.; Guy, A.; Biggin, P. C. *Biophys. J.* **2008**, *95*, 5627–5636.
[235] MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *Biophys. J.* **2008**, *94*, 3393–3404.
[236] Yoo, J.; Cui, Q. *Biophys. J.* **2008**, *94*, L61–L63.
[237] Li, L.; Vorobyov, I.; MacKerell, A. D.; Allen, T. W. *Biophys. J.* **2008**, *94*, L11–L13.
[238] MacKerell, A. D. *J. Phys. Chem.* **1995**, *99*, 1846–1855.
[239] Salaniwal, S.; Cui, S. T.; Cochran, H. D.; Cummings, P. T. *Langmuir* **2001**, *17*, 1773–1783.
[240] Jakobtorweihen, S.; Ingram, T.; Smirnova, I. *J. Comput. Chem.* **2013**, *34*, 1332–1340.
[241] Chaimovich, H.; Aleixo, R.; Cuccovia, I.; Zanette, D.; Quina, F. *Solution Behavior of Surfactants - Theoretical and Applied Aspects*; Plenum Press, N.Y., 1982; Vol. 2; pp 949–974.
[242] de Castro, B.; Gameiro, P.; Lima, J. L. F. C.; Matos, C.; Reis, S. *Lipids* **2001**, *36*, 89–96.
[243] Ingram, T.; Mehling, T.; Smirnova, I. *Chem. Eng. J.* **2013**, *218*, 204–213.
[244] Frisch, M. J. et al. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
[245] Hegazy, L.; Richards, N. G. J. *J. Mol. Modeling* **2013**, *19*, 5075–5087.

[246] Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117–1157.
[247] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
[248] Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
[249] Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845–854.
[250] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
[251] Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.
[252] Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
[253] Perdew, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824.
[254] Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.
[255] Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.
[256] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652–660.
[257] HyperChem. HyperChem Profesional Overview; http://www. hyper.com/?TabId=361 (accessed 29 May 2013).
[258] Linke, D. *Methods in Enzymology*; Elsevier, 2009; Vol. 463; pp 603–617.
[259] Tiller, G. E.; Mueller, T. J.; Dockter, M. E.; Struve, W. G. *Anal. Biochem.* **1984**, *141*, 262–266.
[260] Panayotova, S.; Bivas, I. *Bulg. J. Phys.* **2004**, *31*, 83–86.
[261] Tani, H.; Kamidate, T.; Watanabe, H. *Analyt. Sci.* **1998**, *14*, 875–888.
[262] Brown, W.; Pu, Z.; Rymden, R. *J. Phys. Chem.* **1988**, *92*, 6086–6094.
[263] Magid, L. J.; Triolo, R.; Johnson, J. S. *J. Phys. Chem.* **1984**, *88*, 5730–5734.
[264] Wolszczak, M.; Miller, J. *J. Photochem. Photobiol. A* **2002**, *147*, 45–54.
[265] McCarroll, M. E.; Toerne, K.; von Wandruszka, R. *Langmuir* **1998**, *14*, 6096–6100.
[266] Patel, V.; Ray, D.; Aswal, V. K.; Bahadur, P. *Colloids Surf., A* **2014**, *450*, 106–114.
[267] LeBard, D. N.; Levine, B. G.; Mertmann, P.; Barr, S. A.; Jusufi, A.; Sanders, S.; Klein, M. L.; Panagiotopoulos, A. Z. *Soft Matter* **2012**, *8*, 2385–2397.
[268] Mehling, T.; Ingram, T.; Smirnova, I. *Langmuir* **2012**, *28*, 118–124.
[269] Edwards, D. A.; Luthy, R. G.; Liu, Z. *Environ. Sci. Technol.* **1991**, *25*, 127–133.

[270] Yordanova, D.; Ritter, E.; Gerlach, T.; Jensen, J.-H.; Smirnova, I.; Jakobtorweihen, S. *J. Phys. Chem. B* **2017**, *121*, 5794–5809.
[271] Marina, M. L.; García, M. A. In *Encyclopedia of Separation Science*; Cooke, M., Poole, C. F., Eds.; Academic Press Ltd, England, 2000; pp 726–737.
[272] Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1-2*, 19–25.
[273] Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. *J. Chem. Theory Comput.* **2010**, *6*, 459–466.
[274] Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
[275] Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. *J. Comput. Chem.* **2009**, *30*, 2157–2164.
[276] Torrie, G. M.; Valleau, J. P. *J. Chem. Phys.* **1977**, *66*, 1402.
[277] Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
[278] TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
[279] Diedenhofen, M.; Klamt, A. COSMOconf; Version 3.0; COSMOlogic GmbH & Co. KG: Leverkusen, Germany, 2013.
[280] Eckert, F.; Klamt, A. *AIChE J.* **2002**, *48*, 369–385.
[281] COSMOtherm, Version C3.0, Release 15.01; COSMOlogic GmbH & Co. KG, http://www.cosmologic.de.
[282] Marangoni, D.; Kwak, J. *Solubilization in Surfactant Aggregates*; Marcel Dekker: New York, 1995.
[283] Kelly, K. A.; Burns, S. T.; Khaledi, M. G. *Anal. Chem.* **2001**, *73*, 6057–6062.
[284] Abraham, M. H.; Chadha, H. S.; Dixon, J. P.; Rafols, C.; Treiner, C. *J. Chem. Soc.* **1995**, 887.
[285] Berthod, A.; Garcia-Alvarez-Coque, C. *Micellar liquid chromatography*; Chromatographic science series 83; Marcel Dekker: New York, 2000.
[286] Loureiro, D. R.; Soares, J. X.; Lopes, D.; Macedo, T.; Yordanova, D.; Jakobtorweihen, S.; Nunes, C.; Reis, S.; Pinto, M. M.; Afonso, C. M. *Eur. J. Pharm. Sci.* **2018**, *115*, 369–380.
[287] Paloncýová, M.; DeVane, R.; Murch, B.; Berka, K.; Otyepka, M. *J. Phys. Chem. B* **2014**, *118*, 1030–1039.
[288] Eckert, F.; Klamt, A. COSMOtherm, Version C3.0, Release 15.01. 2014.
[289] Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1085–1100.
[290] Endo, S.; Escher, B. I.; Goss, K.-U. *Environ. Sci. Technol.* **2011**, *45*, 5912–5921.

[291] Rich, P. R.; Harper, R. *FEBS Letters* **1990**, *269*, 139–144.
[292] Abraham, M. H.; Zissimos, A. M.; Huddleston, J. G.; Willauer, H. D.; Rogers, R. D.; Acree, W. E. *Ind. Eng. Chem. Res.* **2003**, *42*, 413–418.
[293] Jämbeck, J. P. M.; Lyubartsev, A. P. *J. Phys. Chem. B* **2012**, *116*, 3164–3179.
[294] Baker, C. M. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 241–254.
[295] Chowdhary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D.; Roux, B. *J. Phys. Chem. B* **2013**, *117*, 9142–9160.
[296] Yordanova, D.; Ritter, E.; Smirnova, I.; Jakobtorweihen, S. *Langmuir* **2017**, *33*, 12306–12316.
[297] Mukerjee, P. In *Micellization, Solubilization and Microemulsions*; Mittal, K. L., Ed.; New York Plenum Press, 1977; Vol. 1; p 171.
[298] Kamayama, K.; Takagi, T. *J. Colloid Interface Sci.* **1990**, *140*, 517.
[299] Tokiwa, F.; Aigami, K. *Kolloid Z. Z. Polym.* **1970**, *239*, 687–691.
[300] Dubin, P.; Principi, J.; Smith, B.; Fallon, M. *J. Colloid Interface Sci.* **1989**, *127*, 558–565.
[301] Komaromy-Hiller, G.; Calkins, N.; von Wandruszka, R. *Langmuir* **1996**, *12*, 916–920.
[302] Dean, J. A., Lange, N. A., Eds. *Lange's handbook of chemistry*, 15th ed.; McGraw-Hill handbooks; McGraw-Hill: New York, NY, 1999; OCLC: 833364385.
[303] Wang, J.; Hou, T. *J. Chem. Inf. Model.* **2010**, *50*, 55–67.
[304] Taylor, R. D.; MacCoss, M.; Lawson, A. D. G. *J. Med. Chem.* **2014**, *57*, 5845–5859.